

Universidad Carlos III de Madrid

Escuela Politécnica Superior



Doctorado en Ciencia y Tecnología Informática

Departamento de Informática

Diseño de un modelo conceptual multi-dominio para
recomendaciones mediante el filtrado de
información semántica en los medios sociales.

Autor: **D. Daniel Villanueva Vasquez**

Directores: **Dr. D. Ángel García Crespo**
Dr. D. Israel González Carrasco

Leganés, Septiembre de 2015



TESIS DOCTORAL

DISEÑO DE UN MOELO CONCEPTUAL MULTI-DOMINIO PARA RECOMENDACIONES MEDIANTE EL FILTRADO DE INFORMACIÓN SEMÁNTICA EN LOS MEDIOS SOCIALES

Autor: Daniel Villanueva Vasquez

Directores: Dr. Ángel García Crespo

Dr. Israel González Carrasco

Firma del Tribunal Calificador:

Firma

Presidente: (Dr. D. Antoni Bibiloni Coll)

Vocal: (Dr. Dña. Belén Ruiz Mezcuca)

Secretario: (Dr. D. Giner Alor Hernández)

Calificación:

Leganés, Septiembre de 2015

It is well with my soul

A Gaby, Jaser, Ximena y Daniela

Agradecimientos

Es muy difícil tratar de resumir en tan pocas líneas la inmensa gratitud que siento hacia tantas personas que me han acompañado hasta este momento.

En primer lugar quiero agradecer a mis directores, Ángel García e Israel González por aguantarme durante estos años, por su infinita paciencia al resolver todas y cada una de mis inquietudes, gracias por transmitir sus conocimientos, su experiencia, pero sobre todo, a su tiempo y dedicación para llevar a cabo la realización de esta tesis doctoral, muchas gracias.

A los profesores, Belén Ruiz Mezcuá, Alejandro Rodríguez González y Juan Miguel Gómez Berbís por contribuir a la predefensa de esta tesis, gracias por sus aportaciones. Asimismo, a Antoni Bibiloni Coll y Giner Alor Hernández por sus comentarios en la defensa de la tesis.

A mis compañeros del grupo IBIS y a mis nuevos amigos que me recibieron y acogieron durante esta investigación, a José Luis, Gandhi, Yuliana, Mateusz, Alejandro, Enrique, Miguel y Javier, gracias por sus consejos.

Gracias a José Luis López Cuadrado por compartir tus conocimientos y por tus valiosas aportaciones para la mejora de este trabajo. A Juanmi, por los grandes contactos y esas fantásticas historias, por tus valiosas palabras de apoyo y motivación durante toda la investigación. A Rubén Posada por tus consejos apropiados para entender las sinergias de trabajo y tus palabras de apoyo.

Gracias al Consejo Nacional de Ciencia y Tecnología (CONACYT) y la Secretaría de Educación Pública (SEP) a través de PROMEP por conceder los recursos económicos para la realización de esta investigación. Asimismo, al proyecto "FLORA: Financial Linked Open Data-based Reasoning and Management for Web Science". (TIN2011-27405).

Y finalmente, a Gaby mi compañera de aventuras, gracias por tu apoyo incondicional en este arduo camino y por ser una mujer excepcional, a Samuel, Ximena y Daniela que son las sonrisas de mi alma, pequeños para entender las cosas de los

grandes, inocentes que aún no lo saben, pero un día lo entenderán que por amor a vosotros una vez más lo pudimos terminar.

En síntesis, agradezco a todas aquellas personas que de una u otra forma han colaborado para la realización de este trabajo y a todos los que me han demostrado su sincero interés, muchas gracias a todos de corazón.

Resumen

Actualmente los usuarios demandan, cada vez, y de manera más intensa, la búsqueda de distintos contenidos almacenados en la Web. Por un lado, existe una gran cantidad de información en la Web y en los medios sociales, esto es debido a la disponibilidad de información sobre los distintos productos, contenidos y servicios que pueden hacer que un usuario se sienta desbordado al intentar discriminar sobre que producto, que contenido o que servicio cubre sus necesidades. Por otro lado, los Sistemas de Recomendación en las distintas áreas de aplicación son cada vez más frecuentes, ya que son útiles para valorar y filtrar esa gran cantidad de información que se encuentra disponible en la Web desde distintos paradigmas. La necesidad de hacer que los procesos de recomendación sean cada vez más claros, que satisfagan y cumplan con las expectativas de los usuarios a supuesto una gran importancia al estudio relacionado con los distintos modelos formales semánticos existentes en los Sistemas de Recomendación aplicados en los medios sociales, además debido a que los usuarios utilizan la Web para publicar, editar y compartir sus contenidos. Por lo tanto, el uso de los distintos modelos formales semánticos para recomendaciones en los medios sociales facilitan la información y, a su vez, aportan un valor añadido al generar una representación del conocimiento sobre distintos dominios, además la información sirve de base para generar recomendaciones a partir de Sistemas de Recomendación a los distintos usuarios en la Web.

La Web semántica posibilita la convergencia entre el uso y la interacción de las personas y los medio sociales, permitiendo crear una gran variedad de contenidos accesibles a las tecnologías semánticas de la Web, a las técnicas de aprendizaje y el filtrado de información. Además, si añadimos que existen las plataformas de comunicación social en la Web, que surgen ante la necesidad de ofrecer una mayor diversidad de información para proporcionar los diversos contenidos personalizados hacia los diferentes tipos de usuarios.

Existen distintos modelos semánticos para Sistemas Basados en Conocimiento que pueden aplicarse en diferentes ámbitos multidisciplinarios, tales como, lenguaje natural, realidad virtual, redes neuronales, juegos masivos, sistemas expertos, robótica, sistemas de planeación, reconocimiento de imágenes, traductores, solución de problemas, sistemas evolutivos y el aprendizaje automático entre otros. Sin embargo, los modelos basados en conocimiento semántico en Sistemas de Recomendación para entornos de medios sociales aún no han sido completamente explotados, constituyendo un área de investigación abierta para la búsqueda de constantes soliviones en la información desde distintos dominios.

Por lo tanto, esta investigación plantea el diseño de un nuevo modelo conceptual multi-dominio semántico para la representación del conocimiento sobre los distintos productos, marcas sus características y servicios ofertados en las redes sociales, a su vez, el modelo conceptual multi-dominio puede modelar y gestionar el conocimiento de diferentes perfiles de usuarios, productos y medios sociales caracterizados para distintos dominios, dentro de un contexto de servicios y productos que, sin cambiar sus conceptos principales, el modelo pueda ser aplicado a distintos dominios para la representación del conocimiento. Además de las hipótesis que marcaron las directrices de trabajo y los objetivos planteados, la presente tesis aporta el diseño del propuesto modelo. La metodología seguida para la elaboración de esta tesis, ha consistido en lo siguiente:

1. – Estudio del estado de la cuestión de la investigación. Dicho análisis permitirá conocer la originalidad y los recursos existentes en el área que se pretende abordar.

2. – Definición de un nuevo modelo conceptual multi-dominio basado en el conocimiento semántico. En paralelo al estudio del estado de la cuestión que permite conocer el estudio del problema y, que a su vez, facilita la definición del modelo. El modelo se desarrollará bajo una herramienta de modelado que facilita la gestión de los conceptos representados en el modelo y, un experto que facilita la interpretación de los datos.

3. – Extracción de datos semánticos basados en contenido estructurado, la información será extraída desde las fuentes de información almacenadas en la Web.

4. – Solución preliminar, dicha etapa nos permite conocer los primeros resultados y un primer comportamiento del modelado a partir de la extracción de datos.

5. – Diseño de un marco computacional. Dicha etapa será el desarrollo de un marco de trabajo basado en el modelo propuesto que integrará un Sistema Basado en Conocimiento, un Sistema de Recomendación, los datos semánticos basados en contenido estructurado semántico y la información que será extraída desde la Web.

6. – Validación y experimentación, en esta fase se ha comprobado las hipótesis planteadas en la investigación, además que el modelo desarrollado puede representar el conocimiento relativo al problema, aplicándolo a la representación del conocimiento para distintos dominios a partir de marco computacional desarrollado y, que a su vez es basado en conocimiento semántico y contenido estructurado.

7. – Verificación y análisis de los resultados. Tras la etapa de validación se estudian los resultados obtenidos que permiten comprobar la validez del modelo propuesto en esta investigación. El objetivo de esta valoración es generar conocimiento para diferentes dominios a partir del modelo conceptual, la información almacenada en el sistema sirve para la generación de recomendaciones a partir de un Sistema de Recomendación. Por último, se presentan las conclusiones extraídas de la etapa verificación y análisis de los resultados que permiten comprobar la validez del modelo y las herramientas propuestas para ésta investigación.

8. – Documentación. A lo largo de todo el proceso de elaboración de la tesis se ha generado la documentación que constituye la presente tesis doctoral.

Las conclusiones del modelo conceptual multi-dominio abre nuevas posibilidades en el área de la Web semántica, Sistemas Basados en Conocimiento y los modelos formales semánticos pertenecientes al área de la Inteligencia Artificial, específicamente en la concepción y desarrollo de un nuevo modelo conceptual multi-dominio. Además, a partir del modelado de técnicas multi-dominio facilita la búsqueda de soluciones en la información, la toma de decisiones y el empleo de conocimiento especializado en diferentes dominios de aplicación de contenido estructurado y semántico, a su vez, generando información relevante sobre los gustos, necesidades y preferencias de los usuarios permitiendo generar recomendaciones a partir de un Sistema de Recomendación.

Índice de contenido

Resumen.....	i
Índice de contenido	v
Índice de figuras.....	xi
Índice de tablas.....	xv
Acrónimos.....	xvii
Capítulo 1.....	1
1. Introducción.....	1
1.1.Contexto	2
1.2.Breve descripción del problema	3
1.3.Motivación.....	4
1.4.Objetivos.....	5
1.4.1. Objetivo general	6
1.4.2. Objetivos específicos	6
1.5.Hipótesis	7
1.6.Justificación	9
1.6.1. ¿Cuál es la importancia de estudiar los modelos para recomendaciones en los medios sociales?	9
1.6.2. ¿Por qué el diseño de un nuevo modelo conceptual multi-domino?	10
1.6.3. ¿Qué aportaciones tendrá esta tesis doctoral?	10
1.7.Metodología de la investigación.....	11
1.8.Estructura del documento	14

Capítulo 2.....	17
2. Estado de la cuestión	17
2.1. Información semántica en los medios sociales	18
2.1.1. La Web semántica	18
2.1.2. Los medios de comunicación social.....	24
2.1.3. La información semántica y los medios sociales	29
2.1.4. Representación del conocimiento en las redes sociales	31
2.1.5. Procesamiento de lenguaje natural.....	35
2.1.6. Ontologías	37
2.2. Clasificación de modelos para Sistemas de Recomendación	42
2.2.1. Modelos de recomendación.....	42
2.2.2. Modelos multi-dominios existentes en los medios sociales	45
2.2.3. Modelos formales semánticos multi-dominio	48
2.3. Sistemas de Recomendación en los medios sociales	62
2.3.1. Introducción	62
2.3.2. Evolución histórica.....	65
2.3.3. Características de modelado de perfiles de usuarios	69
2.3.4. Características de modelado de productos	70
2.3.5. Características de filtrado de información	71
2.3.6. Recomendación basada en contenido.....	73
2.4. Técnicas y aplicaciones para modelos semánticos en medios sociales	77
2.4.1. Modelos para la representación basados en ontologías.....	78
2.4.2. Modelos para la representación en Sistemas Basados en Conocimiento ..	80
2.4.3. Lógica borrosa.....	84
2.4.4. Transformación de componentes y conjuntos difusos	85
2.5. Discusión y conclusión del estado de la cuestión	89
2.6. Sumario.....	91

Capítulo 3.....	93
3. Solución propuesta	93
3.1. Introducción.....	94
3.2. Modelo propuesto	96
3.3. Sumario.....	99
 Capítulo 4.....	 101
4. Modelo conceptual multi-dominio.....	101
4.1. Consideraciones previas	102
4.2. Modelo abstracto	105
4.3. Modelo conceptual multi-dominio	107
4.4. Modelo formal	110
4.4.1. Producto	111
4.4.2. Categoría	112
4.4.3. Perfil.....	113
4.4.4. Categoría de productos.....	113
4.4.5. Características	114
4.4.6. Permanencia	115
4.4.7. Temporal	116
4.4.8. Valoración	117
4.4.9. Clasificación y características de red social	117
4.5. Vistas parciales del modelo conceptual multi-dominio.....	119
4.5.1. Relaciones entre los conceptos principales del modelo conceptual multi- dominio	119
4.5.2. Semántica del modelo	125
4.6. Sumario.....	127

Capítulo 7.....	213
Conclusiones y futuras líneas de investigación	213
7.1. Conclusiones.....	214
7.2. Futuras líneas de investigación.....	216
7.3. Publicaciones realizadas a raíz de esta tesis doctoral	216
 Bibliografía	 221

x

Índice de figuras

Figura 1:1 Etapas a desarrollar durante la investigación.	11
Figura 2:1 Niveles abstractos de representación del conocimiento y, herramientas involucradas en la Web semántica [e.g., Duffy & Bruns, 2006].	21
Figura 2:2 Estructura de capas en la Web semántica [e.g., Berners-Lee, Hendler, & Lassila, 2001].....	22
Figura 2:3 Panorama de los medios sociales.	27
Figura 2:4 Ejemplo de representación esquemática de conjunto de individuos conectados entre sí en una red social.	32
Figura 2:5 Ejemplo de análisis utilizando una herramienta de monitorización en Twitter [e.g., Twittiment].	34
Figura 2:6 Ejemplo de representación esquemática de una ontología.	38
Figura 2:7 Clasificación de técnicas usadas para modelos formales en los medios sociales.	49
Figura 2:8 Relaciones que permiten enriquecen las recomendaciones en los medios sociales.	64
Figura 2:9 Proceso genérico para la generación de recomendaciones en un Sistema de Recomendación.	67
Figura 2:10 Arquitectura de alto nivel para un Sistema de Recomendación basado en contenido según [Valdéz 2012].	74
Figura 2:11 Arquitectura general de un sistema experto [e.g., Lopez-Cuadrado, 2012].	82
Figura 4:1 Modelo abstracto para la representación del conocimiento.	106
Figura 4:2 Modelo conceptual multi-dominio.	108
Figura 4:3 Fragmento de la taxonomía para un criterio denominado <<Producto>>....	111

Figura 4:4 Fragmento de la taxonomía para el concepto <<Categoría>>.	112
Figura 4:5 Fragmento de la taxonomía para un concepto denominado <<Perfil>>. ...	113
Figura 4:6: Fragmento de la taxonomía para un concepto denominado <<Categoría de productos>>.....	114
Figura 4:7 Fragmento de la taxonomía para un concepto denominado características.	115
Figura 4:8 Fragmento de la taxonomía para un concepto de caracterizado <<Permanencia>>.....	116
Figura 4:9 Fragmento de la taxonomía para un concepto de caracterizado <<Temporal>>	116
Figura 4:10 Fragmento de la taxonomía para el concepto de << Valoración>>.	117
Figura 4:11 Fragmento de la taxonomía para la clasificación y características de una red social.....	118
Figura 4:12 Vista parcial abstracta denominada modelo formal.	120
Figura 4:13 Relaciones para el concepto definido como <<Categoría>>.....	122
Figura 4:14 Relaciones y características.....	123
Figura 4:15 Relaciones de alto nivel para una valoración.	124
Figura 5:1 Arquitectura general del marco computacional basada en el modelo multi-dominio.....	133
Figura 5:2 Herramienta de modelado para la gestión de la información	137
Figura 5:3 Definición de criterios y conceptos para la extracción mediante la herramienta de modelado.....	138
Figura 5:4 Proceso genérico para un análisis en la extracción de contenidos.	141
Figura 5:5 Ejemplo del formato CSV para la clasificación del contenido.....	142
Figura 5:6 Modelado de alto nivel para gestión de la información.	144
Figura 5:7 Proceso genérico para la gestión del conocimiento.....	145
Figura 5:8 Pasos en el proceso de recomendación.....	149
Figura 5:9 Herramienta de recomendación.....	154
Figura 5:10 Interfaz de recomendación de productos.....	155

Figura 6:1 Fragmento de la taxonomía en el dominio de la industria automotriz.	165
Figura 6:2 Conceptos y relaciones a partir del modelo formal aplicados al dominio de la industria automotriz.	167
Figura 6:3 Gráfica sobre las menciones que realizan los usuarios hacia los productos del dominio de la industria automotriz en el rango establecido.	169
Figura 6:4 Gráfica de los productos clasificados por marcas.	170
Figura 6:5 Relaciones y características.	172
Figura 6:6 Análisis que menciona relaciones de características.	173
Figura 6:7 (a) Tweets que mencionaba marca Audi, (b) Tweets que mencionan característica de seguridad.	174
Figura 6:8 Relaciones de alto nivel para la validación a partir de un caso de estudio. .	176
Figura 6:9 Menciones de producto y característica	178
Figura 6:10 Palabras claves que mencionan los usuarios.	178
Figura 6:11 Número de menciones para un producto y sus características.	179
Figura 6:12 Análisis de Tweets que mencionan productos de marca Audi, Citroën y Renault.	180
Figura 6:13 Características del dominio en la industria automotriz.	181
Figura 6:14 Tweets que mencionan diferentes características de productos	181
Figura 6:15 Relaciones entre los conceptos del dominio automotriz	183
Figura 6:16 Análisis de características y perfiles	184
Figura 6:17 Usuarios que mencionan distintos productos a través de la red social Twitter	186
Figura 6:18 Usuarios que mencionan distintas características a través de la red social Twitter.	187
Figura 6:19 Conceptos y relaciones a partir del modelo formal aplicados al dominio de la industria de la Innovación y el emprendimiento empresarial.	190
Figura 6:20 Evolución de las menciones de los usuarios sobre productos para el dominio de la industria de la Innovación y el emprendimiento empresarial.	191

Figura 6:21 Evolución de las menciones de los usuarios sobre productos para el dominio de la industria de la Innovación y el emprendimiento empresarial.	192
Figura 6:22 Relaciones de características para el dominio de la innovación y el emprendimiento empresarial.	194
Figura 6:23 Análisis de características mencionadas.....	195
Figura 6:24 Relación de conceptos para generar la validación del caso de estudio	196
Figura 6:25 Análisis de tweets que mencionan distintas características en el dominio empresarial.....	197
Figura 6:26 Fragmento de la taxonomía en el dominio de la industria de la innovación y el emprendimiento empresarial.....	199
Figura 6:27 Herramienta de recomendación, interfaz de acceso a un producto del dominio de la industria automotriz	201
Figura 6:28 Ponderación en el dominio de la industria automotriz	202
Figura 6:29 Relaciones entre características de usuarios y características de productos en el dominio de la industria automotriz	203
Figura 6:30 Lista de características de productos	204
Figura 6:31 Petición de recomendación.....	205
Figura 6:32 Lista de características de productos	205
Figura 6:33 Resultado de recomendaciones en el dominio de la industria automotriz..	206
Figura 6:34 Herramienta de recomendación, interfaz de acceso a un producto del dominio de la industria de la innovación y el emprendimiento empresarial.....	207
Figura 6:35 Relaciones entre características de usuarios y características de productos en el dominio de la industria de la innovación y el emprendimiento empresarial	208
Figura 6:36 Resultado de recomendaciones en el dominio de la industria de la innovación y el emprendimiento empresarial	208

Índice de tablas

Tabla 2:1:	50
Tabla 2:2:	56
Tabla 2:3:	58
Tabla 4:1:	109
Tabla 6:1:	171
Tabla 6:2:	193

Acrónimos

- **SBC.** Sistemas Basados en Conocimiento
- **IA.** Inteligencia Artificial
- **SR.** Sistemas de Recomendación
- **TS.** Tecnologías Semánticas
- **W3C.** World Wide Web Consortium
- **RDF.** Resource Description Framework
- **RDFS.** Resource Description Framework Schema
- **SKOS.** Simple Knowledge Organization System
- **SPARQL.** Protocol and RDF Query Language
- **XML.** Extensible Markup Language
- **OWL.** Web Ontology Language
- **RIF.** Rule Interchange Format
- **DARPA.** Defense Advanced Research Projects Agency (DAML)
- **OIL.** Ontology Inference Layer

- **SRC.** Collaborative Recommender System
- **SRD.** Recommender System Demography
- **SRBC.** System Recommender Content-Based
- **SRKB.** System Recommender Knowledge Based
- **SRH.** System Recommender Hybrid
- **SREFB.** System Recommender Economic Factor Based
- **RSS.** Really Simple Syndication
- **CB.** Content-Based
- **CF.** Filtering-Collaborative
- **KB.** Knowledge-Based
- **URL.** Uniform Resource Locator
- **PACT.** Profile Agregations based on Clustering Transactions
- **CD.** Cross-Domain
- **ARHP.** Association Rule Hypergraph Partitioning
- **KIF.** Knowledge Interchange Format
- **LOOM.** Frame based language
- **OCML.** Operational Conceptual Modelling Language
- **F-Logic.** Frame Logic

- **PLN.** Procesamiento de Lenguaje Natural
- **HTML.** HyperText Markup Language

Capítulo 1

Introducción

El objetivo principal de la presente investigación es la concepción y desarrollo de un nuevo modelo conceptual multi-dominio. El propósito de este capítulo es realizar una breve descripción al problema sobre los modelos formales semánticos basados en el conocimiento. En este sentido proporcionar un modelo multi-dominio para recomendaciones mediante el filtrado de información semántica aplicado a distintos dominios en los medios sociales.

Una vez explicada la motivación de esta investigación y definidos los objetivos, el capítulo detalla las hipótesis que marcan las directrices de trabajo para esta tesis doctoral, a su vez, la justificación sobre la necesidad de un nuevo modelo conceptual. Para finalizar, se detalla la metodología empleada en el desarrollo de la presente tesis doctoral. Además, se realiza un breve resumen de la estructura del documento.

1.1. Contexto

La necesidad de hacer que los procesos de recomendación sean cada vez más claros, que satisfagan y cumplan con las expectativas de los usuarios de una manera sencilla y ágil, ha supuesto que los Sistemas Basados en Conocimiento (SBC) se adentren en el área para la búsqueda de constantes soluciones en la información desde diferentes paradigmas en la investigación de la Inteligencia Artificial (IA). Por lo tanto, las plataformas de comunicación social en la Web surgen ante la necesidad de ofrecer mayor diversidad de información a usuarios que demandan, cada vez y, de manera más intensa, la búsqueda de contenidos masivos almacenados en la Web desde distintos paradigmas. De igual forma, el uso de Sistemas de Recomendación (SR) en diferentes campos de aplicación en la Web es cada vez más frecuente, debido al interés de los usuarios para acceder a múltiples productos potenciales de su beneficio que ofrecen una calidad en los servicios desde diferentes dominios. En los últimos años se ha incrementado la publicación de trabajos relacionados con las áreas de los Sistemas Basados en Conocimiento, Sistemas de Recomendación y las Tecnologías Semánticas (TS), aplicando en ellas distintas técnicas, tales como, aprendizaje automático, modelado de usuario, recuperación de información o la gestión del conocimiento. Todas estas técnicas usadas por dichos sistemas, basan su funcionamiento sobre una base de conocimiento normalmente definida a través de un esquema de conceptos (como una taxonomía o una ontología), para ser aplicados a diferentes dominios.

Los modelos formales semánticos que representan los SBC y los SR contienen una característica en común, la información de un dominio y la utilización de perfiles para representar esa información sobre las necesidades e intereses a largo, mediano o corto plazo de los usuarios. De esta manera, la información de productos y los perfiles de usuario se convierten en una pieza clave para obtener un filtrado de información semántica eficiente, ya que un modelado de perfiles o productos inadecuado, puede conducir a unas recomendaciones de baja calidad y poco relevantes para el usuario. Además, los modelos semánticos son considerados como una rama de la IA aplicada para inferir soluciones a través del conocimiento previo a un contexto de información, sin embargo, no existe un modelo que pueda ser aplicable a distintos dominios sin modificar sus elementos formales

o los conceptos principales de su base original. Por ello, el trabajo que se presenta en la presente tesis doctoral, es el desarrollo de un nuevo modelo conceptual multi-dominio para recomendaciones mediante el filtrado de información semántica aplicado a distintos dominios en los medios sociales. Dentro de esta investigación se desarrolla un marco de trabajo para la demostración y validación de los objetivos planteados al inicio de la tesis doctoral. Dicha investigación supone una aportación en el diseño de los modelos formales semánticos, a la vez, establece una base para los SBC que permiten gestionar el conocimiento que parten de la generación de recomendaciones sobre contenido estructurado y semántico. Además, se resalta la importancia y el uso de los Sistemas de Recomendación en el campo de la IA sobre la utilización de modelos formales para la representación de la información, específicamente su aplicación en los medios sociales. Dicha información será extraída a partir de los medios sociales en Internet.

1.2. Breve descripción del problema

La Web semántica posibilita la convergencia entre el uso y la interacción de las personas y los medio sociales, permitiendo crear una gran variedad de contenidos accesibles a las TS de la Web, a las técnicas de aprendizaje y el filtrado de Información. Por lo tanto, en la Web semántica existe una gran cantidad de información disponible que puede generar reglas para solucionar la problemática asociada al significado, sentido e interpretación de palabras, expresiones, símbolos o representación formal. A su vez, estas reglas aseguran la organización del contenido ya que esa diversidad de datos heterogéneos debe ser procesada de forma rápida, transparente y eficaz.

En este punto de la problemática, los Sistemas Basados en Conocimiento que utilizan técnicas de aprendizaje automático, recuperación de la información y modelos conceptuales son capaces de analizar informes, literatura médica, contenidos multimedia y otras amplias gamas de datos sobre diversos dominios de aplicación. Además, utilizan técnicas de filtrado y recuperación de información semántica para la gestión del conocimiento. Por lo tanto, dichas técnicas permiten la representación de la información y realizan un aprendizaje sobre los datos que son analizados. De esta forma son capaces

de generar y gestionar un contenido que se basa en reglas o conocimiento, las cuales, sin embargo, y en algunas ocasiones, no siempre son correctas y válidas.

No obstante, existen distintos modelos formales semánticos basados en el conocimiento en Sistemas de Recomendación aplicados a los medios sociales. Sin embargo, dichos modelos presentan una desventaja, son modelos que no pueden ser trasladados a un dominio distinto para el que fueron diseñados originalmente, limitándose por tanto, a su uso en áreas específicas y restringidas para los Sistemas de Recomendación.

Debido al interés de generar conocimiento transparente a los usuarios que cada vez demandan con mayor frecuencia la información relativa a los contenidos en la Web. Es importante diseñar modelos conceptuales que permitan representar, procesar y gestionar, desde diferentes perspectivas, un contenido transparente y eficaz a través de la información almacenada en la Web para distintos dominios.

1.3. Motivación

La amplia actividad y presencia de los Sistemas Basados en Conocimiento en diversas áreas y, principalmente en los Sistemas de Recomendación, generan el constante desarrollo de nuevos modelos que ofrecen respuestas claras y precisas para los usuarios, que satisfacen las diversas necesidades de los consumidores y “actores” de una manera sencilla, ágil y eficaz. Según [Andersson et al., 2006], utiliza el término “actores” para distinguir a los usuarios en el modelado.

El uso de las TS y técnicas multi-dominio que se aplican a dichos modelos, permiten explotar el área de las recomendaciones basadas en conocimiento semántico y contenido estructurado. Por lo tanto, surge una necesidad de crear un nuevo modelo que sirva de base para un SBC, un modelo que permita ser llevado a distintos dominios sin cambiar su base de conocimiento original, que facilite una caracterización y contextualización dinámica de las preferencias de los usuarios sobre diversos productos almacenados en los medios sociales y, que a su vez, sirva para generar recomendaciones

basadas en la retroalimentación de los expertos que definen un conocimiento sobre el modelo conceptual a partir de un SR.

En la Web se posibilita la convergencia entre distintas formas de contenido, (Web TV, portales Web de contenidos, protocolo televisión por Internet, televisión digital terrestre, portales de entretenimiento Web, Facebook¹, Twitter², etc.). Esta web considera la información almacenada en las distintas plataformas de comunicación social como un área potencial para las TS que han sido emergentes en la investigación en los últimos años, tales como, la explotación en los modelos aplicados al lenguaje natural, “Data-mining”, aprendizaje automático, basados en conocimiento, agentes expertos y Sistemas de Recomendación. Dichas tecnologías utilizan técnicas de aprendizaje y filtrado de información que permiten la explotación de los modelos conceptuales y, una evolución de los paradigmas de programación para los SBC. Por todo ello, resulta una motivación para el desarrollo de esta investigación.

Por último, un aspecto primordial que se pretende conseguir con esta tesis doctoral es la creación de un nuevo modelo multi-dominio aplicable a distintos dominios sin la necesidad de cambiar sus conceptos principales. Para ello se creará una metodología que guiará las diferentes alternativas disponibles con este fin. Esto supone un desarrollo en la investigación de las TS emergentes, específicamente en los SBC y los modelos formales semánticos.

1.4. Objetivos

Para lograr un objetivo general es preciso definir una serie de tareas encaminadas al diseño de los modelos formales basados en conocimiento semántico y de contenido estructurado, modelos que son frecuentemente usados en los medios sociales y en los SBC. Además, se recabará la información necesaria sobre los distintos modelos formales semánticos existentes en la literatura científica y, se buscará descubrir y evaluar sus

¹ <https://www.facebook.com/>

² <https://twitter.com/>

diferentes áreas de aplicación. Dicha información permitirá respaldar la definición del modelo propuesto que será validado en un marco de trabajo en la investigación. Además, es necesario plantear objetivos específicos para desarrollar tareas concretas que definen la realización de la investigación.

1.4.1. Objetivo general

El objetivo general de esta tesis doctoral es la creación de un modelo conceptual multi-dominio. Dicho modelo servirá como base hacia un Sistema Basado en Conocimiento para representar, a partir de la extracción de información semántica de los medios sociales, información sobre los distintos productos que satisfacen las necesidades de un usuario, a su vez, modelará y gestionará el conocimiento de diferentes perfiles de usuarios, productos y medios sociales caracterizados (contenido caracterizado mediante etiquetas) en múltiples dominios. El modelo permitirá, a partir de un SR, generar recomendaciones desde la información almacenada en las plataformas de comunicación social en la Web mediante la utilización de las TS, técnicas de aprendizaje y procesos de filtrado o recuperación de la información.

1.4.2. Objetivos específicos

De forma desglosada se enumeran los objetivos específicos que permiten alcanzar el objetivo general de la investigación de la tesis doctoral:

- Aplicar las TS, técnicas de aprendizaje y procesos de filtrado de información dentro del campo de la Web semántica para alcanzar la definición de un modelo conceptual, adaptable a cualquier dominio de contenido estructurado y de similitud de contenidos entre distintos usuarios.
- Diseñar un entorno de desarrollo que cumpla con las expectativas del modelo conceptual multi-dominio, que permita aplicar las tecnologías enmarcadas en la

investigación y que, a su vez, establezca una base hacia SBC para múltiples dominios.

- Servir como soporte para la generación de recomendaciones para SR en plataformas que integren a los medios sociales en Internet.
- Gestionar un contenido estructurado y semántico que permita caracterizar productos, perfiles de usuarios y los medios sociales. La caracterización será en función de la similitud entre los contenidos extraídos de los diferentes perfiles y aquellos productos y medios sociales que más se aproximen a las preferencias de los usuarios.
- Caracterizar el contenido estructurado y semántico con similitud entre los medios sociales, perfiles de usuarios y productos.
- Mediante el filtrado de información semántica de los medios sociales y productos, servir de base para generar recomendaciones considerando el contenido caracterizado en el modelo conceptual multi-dominio y las preferencias obtenidas del comportamiento de los usuarios.

1.5. Hipótesis

Gracias al uso de técnicas pertenecientes a la Web es posible generar un nuevo modelo conceptual multi-dominio para Sistemas Basados en Conocimiento, que sirva para representar y generar contenido estructurado y semántico. Este contenido podrá ser utilizado para generar recomendaciones a partir de un Sistema de Recomendación. Para la presente investigación de manera inherente conlleva al surgimiento de ciertas hipótesis, las cuales provienen de una relación entre la teoría y la realidad empírica que enmarca la presente tesis doctoral. En este sentido, una hipótesis sirve para orientar y delimitar una investigación, dándole una dirección definitiva a la búsqueda de la solución de un problema [Tamayo, 2004]. Las hipótesis formuladas a comprobar y, que encapsulan la investigación de la tesis doctoral son las siguientes:

- **H1:** Mediante la capacidad semántica y soporte para contenido estructurado con característica multi-dominio del modelo propuesto es posible la caracterización de productos, medios sociales y Usuarios, para gestionar el conocimiento y ser procesado por SBC a partir de un modelo conceptual multi-dominio. Esta hipótesis lleva implícita la consecución de una serie de hipótesis secundarias que han de cumplirse para validar la hipótesis principal:
 - a) Mediante los medios sociales en la Web es posible caracterizar diferentes productos y perfiles de actores a través del uso de las TS en la Web.
 - b) Mediante la capacidad semántica y de soporte para contenido estructurado del modelo es posible recoger de forma más precisa y exacta la publicación, edición e intercambio de información en los medios sociales.
- **H2:** Es posible ofrecer una solución escalable mediante un nuevo modelo conceptual multi-dominio para recomendaciones en entornos multi-dominios y que sirva como una base para la operabilidad de Sistemas Basados en Conocimiento. Al igual que la hipótesis anterior, será validada a partir de las siguientes hipótesis secundarias:
 - a) A partir de lo anterior es posible la creación de un modelo conceptual multi-dominio que sirva como base de conocimiento para SBC.
 - b) El contenido semántico de los conceptos incluidos en el modelo conceptual permite la creación de una ontología orientada a la generación de recomendaciones personalizadas desde un Sistema de Recomendación.
- **H3:** Es posible diseñar una plataforma que integre el modelo conceptual multi-dominio basado en una solución sobre los contenidos en la Web, medios sociales y usuarios, mediante la anotación semántica y el contenido estructurado.

- **H4:** Es posible la creación de un modelo conceptual multi-dominio, adaptable a cualquier dominio de contenido estructurado y conocimiento semántico.

1.6. Justificación

La representación del conocimiento semántico en las distintas áreas multi-domino, genera, distribuye y enriquece los contenidos personalizados que constantemente son demandados por diferentes géneros de información y distintas audiencias en los medios sociales. Por lo tanto, los modelos conceptuales diseñados para los SBC son un campo muy grande de aplicación y explotación, ya que permiten representar, gestionar y procesar un conocimiento mediante distintas técnicas de filtrado de información en el área de la Web semántica.

1.6.1. ¿Cuál es la importancia de estudiar los modelos para recomendaciones en los medios sociales?

Actualmente existe una gran cantidad de información en los medios sociales, debido a la disponibilidad de información sobre productos, contenidos y servicios puede hacer que un usuario se sienta desbordado al intentar discriminar sobre que producto, contenido y servicio cubre sus necesidades. Por lo tanto, existe una gran importancia del estudio relacionado con los modelos existentes en los Sistemas de Recomendación aplicados en los medios sociales. Además, los usuarios utilizan la Web para publicar, editar y compartir sus contenidos. Por lo tanto, surge la importancia de estudiar los distintos modelos en los SR aplicados a los medios sociales, ya que el uso facilita la información y, que a su vez, aportan un valor añadido al generar una representación del conocimiento sobre diversos dominios que, a su vez, sirve para generar recomendaciones a los distintos usuarios que demandan una información personalizada disponible en la Web.

1.6.2. ¿Por qué el diseño de un nuevo modelo conceptual multi-domino?

Existen distintos modelos semánticos para Sistemas Basados en Conocimiento que pueden aplicarse en diferentes dominios, tales como, lenguaje natural, realidad virtual, redes neuronales, juegos masivos, sistemas expertos, robótica, sistemas de planeación, reconocimiento de imágenes, traductores, solución de problemas, sistemas evolutivos y el aprendizaje automático entre otros. Sin embargo, los modelos basados en conocimiento semántico en Sistemas de Recomendación para entornos de medios sociales aún no han sido completamente explotados, constituyendo un área de investigación abierta para la búsqueda de soluciones en la información desde distintos paradigmas. Por lo tanto, el modelo que se presenta en esta investigación es un modelo multi-dominio que, sin cambiar sus conceptos principales, puede ser aplicado a distintos dominios en los medios sociales para una representación del conocimiento.

1.6.3. ¿Qué aportaciones tendrá esta tesis doctoral?

Esta investigación abre nuevas posibilidades en el área de la Web semántica y los modelos formales semánticos pertenecientes al área de la IA, específicamente en la concepción y desarrollo de modelos ontológicos en los SBC y en la Ingeniería del conocimiento. El modelado de técnicas multi-dominio facilita la búsqueda de soluciones, la toma de decisiones y el empleo de conocimiento especializado en diferentes dominios de aplicación de contenido estructurado y semántico, a su vez generando información relevante sobre los gustos y preferencias de los actores permitiendo generar recomendaciones a partir de un Sistema de Recomendación.

1.7. Metodología de la investigación

En esta sección se detallan las actividades necesarias para alcanzar los objetivos que enmarcan esta investigación descritos en Sección 1.4. Las etapas a desarrollar en las que se ha dividido la presente tesis doctoral son mostradas en la Figura 1:1.

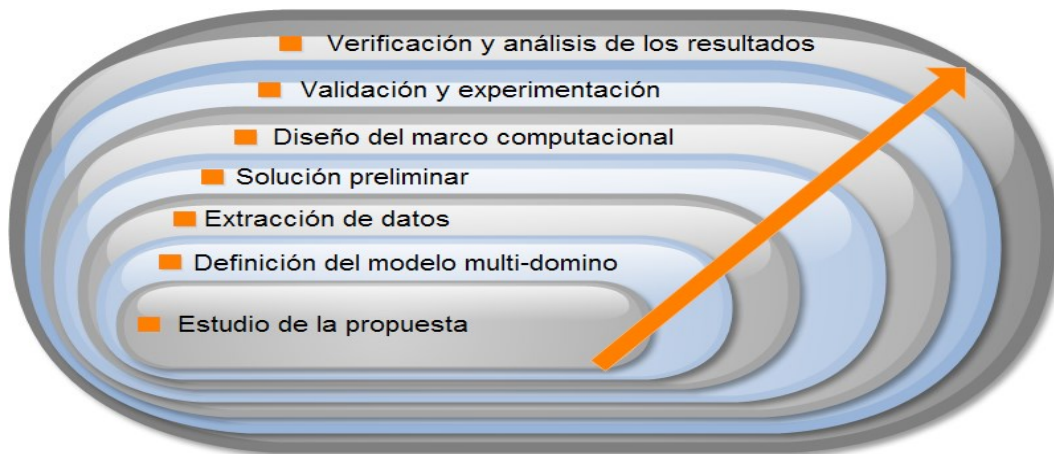


Figura 1:1 Etapas a desarrollar durante la investigación.

Por lo tanto, a continuación se definen las etapas necesarias para adaptarse y conseguir el diseño de un modelo conceptual multi-dominio, definiendo de esta forma una metodología que permite alcanzar los objetivos planteados al comienzo de la investigación.

1. – Una primera etapa es el estudio de la propuesta junto con la literatura del estado de la cuestión de la investigación. Dicho análisis permitirá conocer la originalidad y los recursos existentes en el área que se pretende abordar. Por lo tanto, se ha iniciado con el estudio de diversos trabajos sobre modelos multi-dominio usados en SBC y su aplicación en relación al campo de la IA, área principal de explotación para la presente tesis doctoral.

Como resultado de este estudio, se han identificado los métodos y aplicaciones que se exponen en el Capítulo 2. El estudio se divide en los siguientes pasos:

- Estudio de la información semántica en los medios sociales.
- Estudio del filtrado y recuperación de información semántica en los medios sociales.
- Estudio de Sistemas de Recomendación en los medios sociales.
- Estudio de clasificación y técnicas de modelos para Sistemas de Recomendación.
- Estudio de los Sistemas Basados en Conocimiento basados en modelos de representación.
- Estudio de técnicas y aplicaciones para modelos formales semánticos en los medios sociales.
- Estudio de los distintos modelos semánticos multi-dominios.

2. – La segunda etapa es la definición de un nuevo modelo conceptual multi-dominio basado en el conocimiento semántico. En paralelo al estudio del estado de la cuestión que permite conocer el estudio del problema y, que a su vez, facilita la definición del modelo. El estudio del problema se ha dividido en los siguientes pasos:

- Análisis del modelo conceptual multi-dominio.
- Identificación de la información en los medios sociales.
- Desarrollo del modelo conceptual multi-dominio.
- Ontología del modelo para la gestión y representación del conocimiento.

3. – Una tercer etapa es la extracción de datos semánticos basados en contenido estructurado, la información será extraída desde la Web.

4. – En consecuencia, una tercera etapa es generada para la obtención de una solución preliminar, dicha etapa nos permite conocer los primeros resultados y un primer comportamiento del modelado.

5. – Por lo tanto, la generación de una cuarta etapa que es, el diseño de una plataforma. Dicha etapa será el desarrollo de un marco de trabajo que integrará un SBC, un SR, los datos semánticos basados en contenido estructurado y semántico, información que será extraída desde la Web. El modelo se desarrollará bajo una herramienta de modelado que facilita la gestión de los conceptos representados en el modelo y, un experto que facilita la interpretación de los datos.

6. – En la etapa de validación y experimentación se ha comprobado que el modelo desarrollado pueda representar el conocimiento relativo al problema, aplicándolo a la creación del marco desarrollado y, que a su vez es basado en conocimiento semántico y contenido estructurado. Dicho marco, es capaz de modelar y representar el contenido semántico, información que es extraída de los medios sociales permitiendo la generación de recomendaciones multi-dominio mediante un SR.

7. – Por último, en la etapa de verificación y análisis de los resultados y tras la etapa de validación se estudian los resultados obtenidos que permiten comprobar la validez de la investigación. El objetivo de esta valoración es generar conocimiento referido a la retroalimentación del SR. Generando de esta forma un histórico del sistema que mejora sus recomendaciones a medida que el usuario valora las preferencias previas. Por último se presentan las conclusiones extraídas de la etapa verificación y análisis de los resultados.

En consecuencia a cada una de las etapas propuestas, se llevará a cabo la tarea de documentación y conclusiones que se cumplirán a lo largo de todo el proceso de

Investigación en la presente tesis doctoral. En las primeras etapas se registran los resultados del estudio del arte y del análisis sobre el planteamiento del problema. Durante el desarrollo se describen los diferentes conceptos del modelo multi-dominio, sus aspectos relevantes para el modelado y el marco de desarrollo que integra los componentes definidos y los medios sociales. En la etapa de validación se han documentado los resultados de los distintos procedimientos descritos en los capítulos anteriores.

El cumplimiento de estas etapas permite plantear una serie de actividades a llevar a cabo durante el desarrollo de la presente tesis doctoral. A continuación se desglosan los capítulos que estructuran el documento.

1.8. Estructura del documento

Para una mejor comprensión de la investigación realizada, este documento ha sido estructurado en una serie de capítulos, los cuales se explican a continuación de forma resumida.

- **Capítulo 2. Estado de la cuestión** – En este capítulo se revisan los trabajos realizados hasta la fecha en el área de los Sistemas de Recomendación, los modelos semánticos. Además, se repasan los aspectos relativos a la representación del conocimiento y el filtrado de información semántica en los medios sociales.
- **Capítulo 3. Solución propuesta** – En este capítulo se exponen las razones que justifican el diseño del modelo conceptual multi-dominio para recomendaciones mediante el filtrado de información semántica propuesto en esta investigación.
- **Capítulo 4. Modelo conceptual multi-dominio** – En este capítulo se describe de forma detallada las definiciones abstractas sobre el término de modelo conceptual y el punto hacia donde se pretende abordar la solución del problema en la investigación.

- **Capítulo 5. Arquitectura de la Solución propuesta** – En este capítulo se describe la representación del proceso de extracción de la información semántica de los medios sociales, empleando el modelo propuesto.
- **Capítulo 6. Validación** – En este capítulo se describe la validación del modelo propuesto y su aplicación al dominio de los medios sociales para recomendaciones a partir de dos casos de estudio que permiten validar el modelo y la comprobación de las hipótesis planteadas en esta investigación.
- **Capítulo 7. Conclusiones y futuras líneas de investigación** – En este capítulo se plantean las conclusiones de la investigación que se ha llevado a cabo y se apuntan las líneas en las que se pueden continuar el trabajo realizado. Además, se enlistan las publicaciones realizadas en diferentes ámbitos como consecuencia de la presente investigación.

Capítulo 2

Estado de la cuestión

Actualmente, en un entorno donde la información está accesible en la Web, los SBC se desplazan al desarrollo de modelos formales para los entornos de medios sociales. Por lo tanto, la creación de un nuevo modelo conceptual multi-dominio permitirá la personalización y caracterización de productos, perfiles y medios sociales para diferentes dominios, el modelo utilizará el contenido de información ofrecida en las redes sociales y servirá como base para un SBC. Al mismo tiempo, las preferencias, gustos y necesidades de los actores servirán para generar recomendaciones a partir de un Sistema de Recomendación. En consecuencia y reafirmando el objetivo que se pretende alcanzar durante esta investigación es: ofrecer un modelo para la representación del conocimiento que no se limite a un único dominio, un modelo que pueda ser llevado a distintos dominios sin cambiar sus conceptos principales en su base de conocimiento. Además, desde la caracterización de contenido, dicha información servirá para la generación de recomendaciones sobre los diversos productos que existen en los medios sociales.

La revisión presentada en este capítulo trata de cubrir la idoneidad de los modelos que existen en los SBC y una revisión del estado de la cuestión ubicando en un contexto a la investigación sobre los distintos modelos semánticos existentes en los SR aplicados en los medios sociales, área que se pretende abarcar con la extensión de la información sobre el tema. El análisis de dichos modelos permite describir la comparación de los modelos formales semánticos existentes en los medios sociales, además detallar sus aplicaciones entre los diferentes dominios. Finalmente, conocer las distintas técnicas y aplicaciones para distintos dominios en los medios sociales.

2.1. Información semántica en los medios sociales

2.1.1. La Web semántica

La Web semántica es una Web extendida, dotada de mayor significado, en la que cualquier usuario de Internet puede encontrar respuestas a sus preguntas de forma más rápida y sencilla gracias a una información mejor definida [Zhong, Liu, & Yao, 2002]. Erróneamente se identifican los conceptos de Web semántica y Web 3.0 [Cardoso, 2007]. De manera que, al aclarar la diferencia, se entiende que la Web semántica es un conjunto de actividades, tal y como indica el propio consorcio de la W3C³ (*World Wide Web Consortium*), una comunidad que desarrolla tecnologías para el crecimiento de la Web. Por lo tanto, dichas tecnologías son desarrolladas en conjunto para su aplicación en muchos ámbitos, tales como, redes sociales, publicación de datos, realización de inferencias, marcado semántico de documentos convencionales, etc. Por otra parte, la Web 3.0 se refiere a un entorno en el que aplicaciones y agentes de usuario intercambian datos, los procesan e incluso realizan procesos de inferencias para generar nueva información [J. Hendler, 2009].

Actualmente, la Web es un espacio preparado para el intercambio de información diseñado para el consumo humano en una Web extendida, dotada de un mayor significado en la que cualquier usuario en Internet podrá encontrar respuestas a sus preguntas de forma más rápida y sencilla gracias a una información mejor definida. Al dotar a la Web de más significado y, por lo tanto, de más semántica, se pueden obtener soluciones a problemas habituales en la búsqueda de información gracias a la utilización de una infraestructura común, mediante la cual, es posible compartir, procesar y transferir información de forma sencilla. Esta Web extendida y basada en el significado, se apoya en lenguajes universales que resuelven los problemas ocasionados por una Web carente

³ W3C: <http://www.w3c.es/>

de semántica en la que, en ocasiones, el acceso a la información se convierte en una tarea difícil y frustrante.

La Web semántica se ha descrito en bastantes y de diferentes formas. Por ejemplo, como una visión utópica [Baumann, 2009], como una red de datos [Fundulaki et al., 2012] o simplemente como un cambio de paradigma natural en nuestro uso diario de acceso a la Web [Fensel, 2004]. No obstante y por encima de todo, la Web semántica ha inspirado y, a muchas personas para crear tecnologías y aplicaciones semánticas innovadoras que en los último años el crecimiento de la Web tradicional ha sido constante [Shadbolt, Hall, & Berners-Lee, 2006]. Por un lado, ha permitido dotar a la Web de más significado y, por lo tanto, de más semántica, de manera que se pueden obtener soluciones a problemas habituales en la búsqueda de la información desde diferentes paradigmas [Alba, 2007].

Por otro lado, Lassila & Harmelen comentan que su utilización es más específicamente en la colección, estructuración y recuperación de datos vinculados [Lassila & Harmelen, 2000]. Además, están habilitados por formatos y tecnologías, que a su vez, proporcionan una descripción formal de conceptos, términos y relaciones dentro de un ámbito de conocimiento conocido como estándares del W3C, los cuales incluyen:

- **Resource Description Framework (RDF).** – Es un método general para describir la información.
- **RDF Schema (RDFS).** – Es un conjunto de clases con ciertas propiedades utilizando el lenguaje extensible de representación del conocimiento RDF.
- **Simple Knowledge Organization System (SKOS).** – Es un diseño recomendado por el consorcio W3C para la representación de esquemas de clasificación, taxonomías o cualquier otro tipo de vocabulario estructurado.
- **SPARQL,** acrónimo del inglés (Protocol and RDF Query Language). – Es un lenguaje de consultas, es decir, un lenguaje de consulta para bases de

datos, capaz de recuperar y manipular datos almacenados en formato Resource Description Framework RDF.

- **XML** (Extensible Markup Language). – Es un lenguaje de etiquetado extensible, su función principal es describir datos. Además, se puede definir como un metalenguaje con el que se puede generar otros lenguajes, cuya gramática puede definirse gracias a sus etiquetas que intentan estructurar la información de forma abstracta.
- **Notation3 (N3)**. – Es un formato diseñado con legibilidad en mente, es mucho más compacto y fácil de leer que la notación XML y RDF.
- **Turtle**. – Es un formato para expresar datos y modelo de datos con la sintaxis similar a SPARQL. RDF, a su vez, representa la información utilizando "triples", cada uno de los cuales consta de un sujeto, un predicado, y un objeto, es un formato para almacenar y transmitir los datos.
- **Web Ontology Language (OWL)**. – Es una familia de lenguajes de representación del conocimiento o lenguajes de ontologías para la edición de ontologías o bases de conocimiento.
- **Rule Interchange Format (RIF)**. – Es un diseño de recomendación del consorcio W3C, es parte de la infraestructura para la web semántica, junto con SPARQL, RDF y OWL.

La representación del conocimiento en la Web semántica se utiliza para describir los aspectos de contenido y recursos formales. Este aspecto de la representación de los recursos web que se describe en RDF, la capa más evidente, pero dentro de este aspecto también hay otros niveles menos visibles. En cada uno de los niveles utilizan herramientas

de aplicación sobre la base de esquemas conceptuales y herramientas lógicas. La Figura 2:1, muestra los niveles abstractos de las lógicas, las herramientas y los niveles de representación del conocimiento involucrado en la Web semántica [Bosch, 2006].

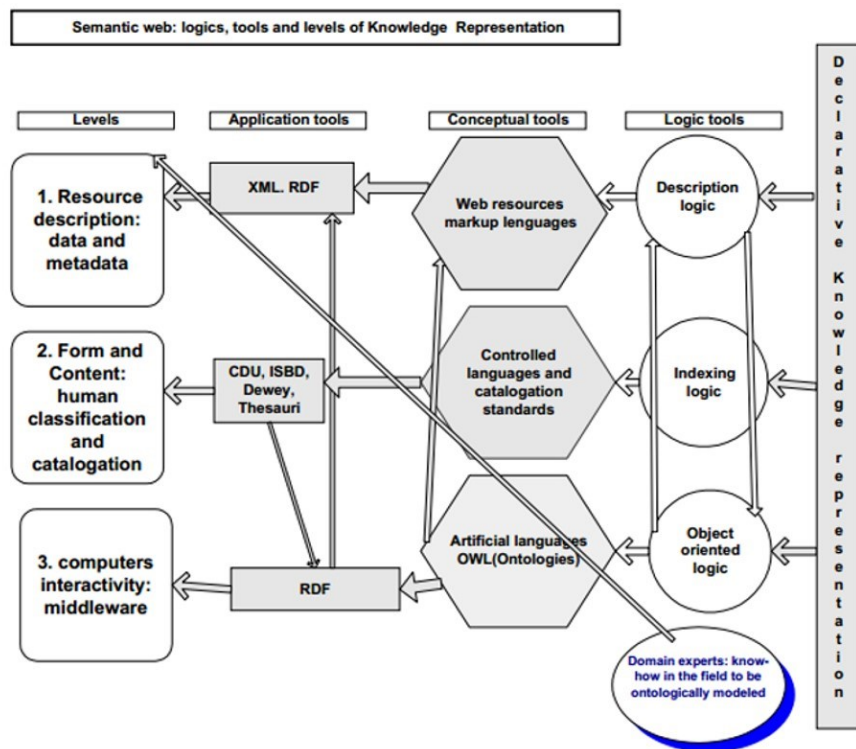


Figura 2:1 Niveles abstractos de representación del conocimiento y herramientas involucradas en la Web semántica [e.g., Duffy & Bruns, 2006].

Debido a los estándares de la Web, es posible establecer normas que garantizan la interoperabilidad en los distintos niveles. Además, la información es procesable a los servicios inteligentes tales como, agentes de información, agentes de búsqueda, y filtros que ofrecen una información de mayor funcionalidad e interoperabilidad en los servicios [Decker & Melnik, 2000]. Actualmente el proceso de construcción de la Web semántica es un área de alta actividad. Su estructura tiene que ser bien definida para realizar las distintas tareas. Por lo tanto, se debe comenzar con las tareas más simples. Los siguientes pasos nos muestran la dirección hacia donde se dirige la Web semántica.

1. Proporcionar una sintaxis común para los estados comprensibles de una máquina.
2. El establecimiento de vocabularios comunes.
3. Convenir en un lenguaje lógico.
4. Usando el lenguaje de intercambio de pruebas.

Berners-Lee propuso una estructura de capas para la Web semántica. Esta estructura refleja los pasos indicados anteriormente. De manera que, sigue el entendimiento de que cada paso por sí solo ya aportará valor añadido, por lo que la Web semántica se puede realizar de forma incremental [Stume et al., 2006]. La Figura 2:2 muestra una visualización de las distintas tecnologías usadas para construir la Web semántica a modo de capas según el nivel de abstracción de cada una de éstas.

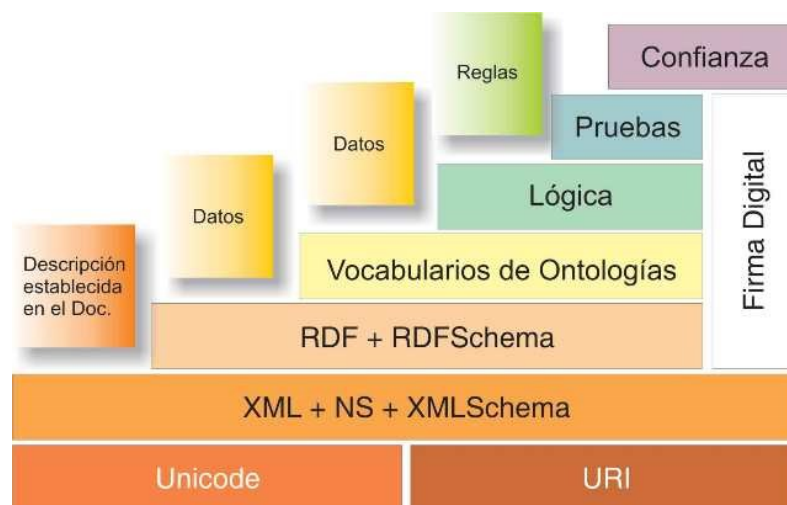


Figura 2:2 Estructura de capas en la Web semántica [e.g., Berners-Lee, Hendler, & Lassila, 2001].

La Web semántica se basa en una visión de Tim Berners-Lee que sugiere enriquecer la Web con información procesable por una máquina, que a su vez, ayuda al usuario en la búsqueda de soluciones a sus distintas tareas. Por ejemplo, los motores de

búsqueda, que en la actualidad son herramientas bastante potentes. Por lo tanto, es posible mejorar la precisión y recuperación con una información procesable por una máquina, ya que puede señalar mediante el motor de búsqueda a las páginas pertinentes [Fensel, 2005].

Por otro lado, el uso de las TS para la representación del conocimiento ha causado una dramática transformación en la práctica de la investigación en diversos dominios [Maedche, Motik, & Stojanovic, 2003]. En consecuencia, un rápido crecimiento en la investigación y desarrollo en la utilización de dichas tecnologías en los sistemas de información, que a su vez, han desatado una atención a nivel mundial por parte de los usuarios en la administración y manejo del conocimiento. Las TS no solamente son útiles en el campo para el que fueron diseñadas, también son aplicables a múltiples entornos y dominios. Por ejemplo, Gandon & Sadeh presentan un trabajo mediante tecnologías semánticas destinado a apoyar la identificación automática y el acceso de los recursos personales en el dominio de la educación [Gandon & Sadeh, 2004], del mismo modo, Frew & Bose presentan una implementación para la gestión de datos en el dominio de las comunicaciones [Frew & Bose 2001]. Asimismo, las TS proporcionan una base consistente y fiable que se puede utilizar para hacer frente a los desafíos relacionados con la organización, manipulación y visualización del conocimiento de datos, además de jugar un papel crucial en el desarrollo de bases tecnológicas para los sistemas en la inteligencia computacional. Estas tecnologías se basan en técnicas convencionales de nuevas y diversas disciplinas dentro de la Ciencias de la Computación, incluyendo la ingeniería del conocimiento, procesamiento del lenguaje natural, inteligencia artificial, bases de datos, los agentes software, etc. Los métodos y herramientas desarrolladas e integradas para este propósito son genéricos y tienen una muy gran potencialidad de aplicación en muchos campos, tales como la recuperación de información, búsquedas semánticas, integración de la información, información sobre interoperabilidad, bioinformática, salud, el aprendizaje, ingeniería de software, el comercio electrónico, la administración electrónica, medios sociales, etc. [Valencia-García, 2014].

El término Tecnologías Semánticas (para hacer referencia a la conocida previamente como Web semántica) ha sido acuñado recientemente para designar una serie de tecnologías, relacionadas principalmente con la Web de nueva generación [R

Studer, Grimm, & Abecker, 2007]. Además, Davies & Weeks, comentan que en la Web los contenidos son algo más que una gran suma de información y servicios escasamente estructurados [Davies & Weeks, 2004]. De manera que este nuevo enfoque propone reestructurar y enriquecer los documentos y componentes en la Web con información semántica explícita, independiente de la presentación que se ofrezca al usuario. Asimismo, susceptible de ser procesada de forma automática por un programa [L. Li & Horrocks, 2004].

En consecuencia, se considera que la Web semántica añadirá la estructura al contenido semántico de los documentos electrónicos, creando un entorno donde los agentes software podrán realizar tareas de manera eficiente [Berners-Lee et al., 2001].

La Web semántica representa la base ideológica y tecnológica de diversos contenidos que son generados por el usuario según [Murugesan, 2007]. Además, la Web puede verse como la suma de distintas formas en que la gente hace uso de los distintos medios de comunicación social para generar diferentes tipos de contenidos [Sizov, 2010].

2.1.2. Los medios de comunicación social

En poco más de una década desde su aparición, la World Wide Web se ha convertido en un instrumento de uso cotidiano en nuestra sociedad, comparable a otros medios tan importantes como la radio, la televisión o el teléfono a los que aventaja en muchos aspectos [Nallasivan & Price, 2012]. La Web es hoy un medio extraordinariamente flexible y económico para la comunicación, el comercio y los negocios, ocio y entretenimiento, acceso a información y servicios, difusión de cultura, etc. Paralelamente al crecimiento espectacular de la Web, las tecnologías que la hacen posible han experimentado una rápida evolución [Felfernig & Friedrich, 2006]. De manera que, la comunicación en los medios sociales se ha intensificado en los últimos años dado el interés por parte de la sociedad [Zeng, Chen, Lusch, & Li, 2010].

El ser humano desde siempre se ha comunicado a través de medios sociales cuales fuera que sean [Leung et al., 2013]. Por ejemplo, la familia, los compañeros del

colegio, los amigos de la universidad, los colegas de profesión por comentar algunos de ellos, todos ellos son redes sociales dentro del contexto de una comunicación social.

Los medios de comunicación social o simplemente medios sociales, actualmente se han convertido en una comunicación basada en la sencillez y la interlocución que hacen posible tener una relación “de tú a tú”. Asimismo, se definen por la convergencia de individuos en redes sociales, son plataformas participativas en los que la información es definida como una categoría de comunicación en línea donde las personas editan, crean y comparten numerosos tipos de contenidos mediante el uso de las tecnologías de la Web a un ritmo prodigioso [Gilbert & Karahalios, 2009]. Por otra parte, los medios sociales están cambiando rápidamente la comunicación del público con una sociedad que es cada vez más inteligente y participativa, que a su vez, demanda una información personalizada así como el establecimiento de tendencias sobre distintos temas y dominios de aplicación [Asur & Huberman, 2010]. Por ejemplo, en el dominio de la accesibilidad donde se puede pensar que no son de gran ayuda, resulta que es todo lo contrario, los medios de comunicación social ofrecen de forma específica, información para que las empresas se comunican con las personas con diferente discapacidad sobre la accesibilidad de sus productos y servicios [Brady & Bigham, 2014]. Además, facilitan el dialogo con los consumidores a través de una información activa con un mecanismo de difusión unidireccional [Lovejoy, Waters, & Saxton, 2012].

Por lo tanto, es necesario describir el concepto de medios de comunicación social, los conceptos relacionados con la Web y el contenido generado por los usuarios. Además de obtener una clasificación de los distintos tipos de medios de comunicación social así como conocer su importancia para su utilización en las distintas áreas de aplicación.

Los medios sociales se diferencian de los conceptos relacionados aparentemente intercambiables de la Web y el contenido generado por los usuarios [Leung et al., 2013]. Por lo tanto, tiene sentido comentar de dónde vienen los medios sociales y lo que contienen.

En 1979, Tom Truscott y Jim Ellis de la universidad de Duke habían creado Usenet, un sistema de discusión en todo el mundo que permite a los usuarios de Internet enviar mensajes públicos [Kaplan & Haenlein, 2010]. Sin embargo, la era de los medios

sociales como lo entendemos hoy en día probablemente comenzó unos 20 años antes, cuando Bruce y Susan Abelson fundaron "Open Diary", un sitio de redes sociales en una versión piloto que reunió a escritores del diario en línea en una sola comunidad [Kaplan & Haenlein, 2009]. La mayoría de veces el término de medios sociales es usado para referirse a las actividades que integran la tecnología, las telecomunicaciones móviles y la interacción social, en forma de conversaciones, fotografías, imágenes, vídeos y pistas de audio [Correa, Hinsley, & de Zúñiga, 2010]. De manera que las conversaciones se convierten en un medio social de comunicación donde las personas comparten sus intereses y experiencias con otros, de una forma natural. Cabe señalar que el desarrollo tecnológico en los últimos años ha sido uno de los más importantes en la historia de la sociedad, no solo para cada campo en específico sino porque se ha universalizado, y ha terminado por afectarnos a todos los usuarios en las múltiples facetas de la vida [Kaplan & Haenlein, 2010].

Los medios sociales de comunicación son distintos de los medios de comunicación industrial, tales como periódicos, canales de televisión y emisoras de radio [Decotignie & Pleinevaux, 1993]. Una de las diferencias entre los medios sociales y los medios de comunicación industrial es, por una lado los medios sociales usan herramientas relativamente baratas que permiten a cualquier persona, publicar y tener acceso al contenido, mientras que los medios de comunicación industriales, por otro lado y por lo general, requieren un capital financiero para iniciar operaciones, tales como máquinas sofisticadas para la impresión, equipos y antenas para una emisora de radio o la concesión de una licencia del espectro radioeléctrico [Veil, Buehner, & Palenchar, 2011]. No obstante, ambos conceptos comparten una característica en común, la capacidad de llegar a un público grande para suplir una necesidad aunque decidan dedicarse a un pequeño nicho; por ejemplo, una publicación en un programa de TV de un medio tradicional pueden llegar a millones de personas en muchas partes del mundo.

Kaplan y Haenlein definen a los medios sociales como: un grupo de aplicaciones basadas en Internet que se desarrollan sobre los fundamentos ideológicos y tecnológicos de la Web 2.0, y que permiten la creación y el intercambio de contenidos generados por el usuario [Kaplan & Haenlein, 2011]. Por lo tanto, los medios sociales llevan adherida una nueva forma de entender la comunicación mediante la colaboración, la instantaneidad de

conversación y la edición del contenido. Dicha información es compartida por los usuarios a través de un conjunto de tecnologías y herramientas en Internet que propician un trabajo creativo y colaborativo en la Web.

Además de los usuarios que utilizan los medios sociales para crear, modificar, compartir y discutir el contenido almacenado en la Web, empresas y ejecutivos también hacen uso de los medios para incrementar sus ventas y ofrecer nuevos productos con un alto impacto a las necesidades de la sociedad que demanda a los usuarios [Kietzmann et al., 2011]. Por lo tanto, el rápido desarrollo de las plataformas de comunicación ha reformado la manera en que las personas y empresas interactúan entre sí, cambiando sus estrategias de negocio y herramientas para comunicarse con sus clientes, al mismo tiempo, generando nuevas oportunidades de crecimiento para las empresas.

La Figura 2:3, muestra un panorama de los medios sociales y las nuevas formas de comunicación entre las empresas y sus clientes, la nueva forma de crear una plataforma de red social organizada, lo que permite la creación de contenidos o la exigencia de una comunicación directa con los consumidores. La nueva forma de comunicación hace que sus seguidores se adentran en el mundo de los medios sociales.



Figura 2:3 Panorama de los medios sociales.

Los medios sociales pueden adoptar distintas formas, tales como, de comunicación, colaboración, multimedia, entretenimiento que permiten a los actores ser activos o pasivos en la generación de contenido en la Web [Cromity, 2012]. Una clasificación de los medios sociales on-line de forma genérica se pueden dividir en tres tipos:

- Medios sociales para publicar
- Medios sociales para compartir
- Medios sociales para conectar

Por lo tanto, los medios sociales son de gran importancia a la hora de influir sobre todo tipo de actores y entornos y, poseen además, un público que es cada vez más participativo. Al mismo tiempo, los medios sociales aportan un gran enriquecimiento, ya que proporcionan las herramientas para un nuevo tipo de relaciones sujetas a una sociedad mayormente globalizada y exigente.

Así mismo, los medios de comunicación social en la Web configuran un tipo específico de información de filtrado, articulada, entre otros, mediante blogs, noticias, música, planes de viaje, páginas web, imágenes, etiquetas, comunidades de contenido, proyectos de colaboración, redes sociales, mundos virtuales de juegos, mundos virtuales sociales; todos ellos, susceptibles de interesar a los usuarios [Choi & Han, 2013].

Los medios sociales es una de esas frases que mucha gente podría conocer, ya que es la combinación de dos palabras familiares, pero a su vez, pueden llegar hacer tan complejas para una nueva comunicación entre las personas de forma on-line, son sin duda el resultado de una evolución a las comunicaciones y las relaciones personales que vienen de mano de la revolución tecnológica hoy en día a través de la Web [Weinberg & Pehlivan, 2011]. Por lo tanto, las relaciones entre los usuarios son basadas en comunicaciones más abiertas, colaborativas y bidireccionales [Enders & Hungenberg, 2008]. Teniendo en cuenta las definiciones antes descritas, cabe resumir que los medios sociales son un grupo de aplicaciones basadas en Internet que se fundamentan en los elementos ideológicos y tecnológicos de la Web, además permiten la creación, edición y el intercambio de diferentes contenidos generados por los usuarios.

El creciente desarrollo de los medios sociales y la Web semántica, resultan un área pujante en la confluencia de la IA y TS [U Bojars & Breslin, 2008]. La Web propone introducir descripciones explícitas sobre el significado de los recursos, además de permitir que las propias máquinas tengan un nivel de comprensión suficiente como para hacerse cargo de una parte, ya sea la parte costosa, la parte rutinaria o la parte que es físicamente inabarcable en la labor del trabajo. Para ello, es necesario hacer que la información que reside en la Web semántica sea entendible por las propias máquinas y sirva a los usuarios con eficacia [Berners-Lee & Kagal, 2008; Shah, Finin, & Joshi, 2002].

2.1.3. La información semántica y los medios sociales

Actualmente, la Web semántica como fuente de información es fundamental y se encuentra al alcance de casi todos los usuarios, la amplia información accesible con carácter prácticamente universal, ha llevado al éxito a muchas organizaciones que hacen uso de la información almacenada en la Web. Las aplicaciones prácticas de esta visionaria Web la convierten en un espacio activo de socialización para la mayoría de los usuarios [Kahney, 2003]. En un estudio Jeff Bar responde: “Encontrar información, sin duda. Hay tantas Webs, tanta información disponible que es difícil encontrar lo que a uno realmente le interesa o le es útil. ¡Una sobredosis de información, definitivamente! ésa sobredosis puede convertir en auténticos retos lo que son simples necesidades” [Redondo, 2007]. Por lo tanto, la información semántica es vital para procesar el contenido, razonarlo, combinarlo y realizar deducciones lógicas, atendiendo a las demandas de información de los usuarios [Mika, 2005]. Mangold & Faulds, mencionan a los medios sociales como una “comunicación generada por los usuarios” ya que en la actualidad representan una amplia fuente de información [Mangold & Faulds, 2009].

Por otro lado, la información en los medios sociales son un movimiento popular en la Web semántica, ya que es sencillo hacer relaciones sociales y personales de confianza [Golbeck, 2006]. Los medios sociales son un dominio muy activo y de rápido movimiento, ya que la información que puede ser al día de hoy podría haber desaparecido del paisaje

virtual mañana. Además, la información semántica en los medios es fundamental. Por ejemplo, las empresas hoy en día consideran el uso de los medios sociales como herramientas de comunicación con sus clientes, ya que a través de ellas, pueden dar a conocer sus marcas de negocios, ofrecer sus servicios y ofertas que son demandadas sobre los distintos tipos de dominios, específicamente a través de las redes sociales [Michaelidou, Siamagka, & Christodoulides, 2011].

En efecto, los medios sociales en la actualidad representan una fuente de información semántica, de tal modo que a las empresas les ha llevado a modificar sus estrategias para comunicarse, ya que ahora el control de la información recae sobre el usuario que demanda una información personalizada [Kinsella et al., 2009]. Dicha información está compuesta de un contenido en línea, tales como, blogs, salas de chat, sitios web de calificación, los sitios web de vídeo que permiten compartir fotos y podcasts, que a su vez, pueden ser creados y utilizados con fines educativos sobre temas que incluyen productos y marcas. Por ejemplo, en particular, Facebook uno de los sitios más populares y visitados de la Web en el mundo [Bulte & Wuyts, 2007].

Ejemplos destacados incluyen sitios como Flickr⁴ y YouTube⁵, para compartir fotos y videos, blogs y sistemas de wiki como Blogger, sitios de marcado sociales como Delicious⁶, sitios de redes sociales como MySpace⁷ y Facebook, y los sitios de microblogging como Twitter. Este último es un servicio que permite a los usuarios enviar mensajes de texto sin formato, llamados tweets, de corta duración, con un máximo de 140 caracteres que muestran la información del usuario en una página de inicio que pueda interesar a otros usuarios almacenados en la Web.

En consecuencia, hoy por hoy, desarrolladores trabajan utilizando las tecnologías de la Web semántica para aumentar las formas en que se crean, utilizan y comparten la información semántica en los medios de comunicación social [Breslin & Decker, 2007], específicamente en las redes sociales y la representación del conocimiento [Rodríguez-González, 2012], ya que existen redes sociales en todas partes, a nuestro alrededor, en

⁴ <https://www.flickr.com/>

⁵ <https://www.youtube.com/>

⁶ <https://delicious.com/>

⁷ <https://myspace.com/>

los lugares de trabajo, así como en las familias y los distintos grupos sociales. Por lo tanto, las redes sociales están diseñadas para ayudarnos a trabajar en conjunto sobre las actividades e intereses comunes.

2.1.4. Representación del conocimiento en las redes sociales

En los últimos años la creciente influencia de los medios sociales ha sido desorbitada, ya que permiten la creación de conocimiento mediante las contribuciones de los usuarios simplificada a través de blogs, wikis, y el despliegue de las redes sociales en línea [Kinsella et al., 2009]. Por lo tanto, esto requiere mecanismos de representación de datos y aplicaciones de la Web social en forma interoperable y extensible. No obstante, la Web semántica proporciona tales mecanismos de representación. Por ejemplo, Bojars, Breslin & Decker describen en su obra los métodos para describir personas y objetos que se conectan explícita o implícitamente en las redes sociales, además cómo los desarrolladores utilizan la Web semántica para representar el conocimiento [Bojars, Breslin, & Decker, 2008]. En consecuencia, la representación del conocimiento es un área de la IA cuyo objetivo es la representación del conocimiento de una forma que permita obtener conclusiones a partir de dicho conocimiento [Davis, Shrobe, & Szolovits, 1993]. Además, la representación del conocimiento en las redes sociales tiene como objetivo representar un cierto conocimiento de un dominio a través de los datos e información almacenada en la Web [Michaelidou et al., 2011].

Diferentes términos se han sugerido para designar a este tipo de representación, aunque los términos más aplicados según Richards son: «Diagramas de conceptos, grafos conceptuales, redes semánticas o redes conceptuales, nombres dados a la misma idea, la de representar información de carácter conceptual de forma gráfica» [Richards & Richards, 1994]. De manera que las redes sociales son un instrumento para el análisis de los datos en la representación del conocimiento.

Las redes sociales son aplicaciones que permiten a los usuarios conectarse mediante la creación de perfiles de información personal, se basan en el establecimiento de las relaciones entre individuos, grupos de personas, departamentos organizativos o corporaciones, que conduce a la creación de redes sociales [Wolfe, 1997 ; Robins et al. 2007]. En general, las redes sociales son beneficiosas para los participantes de las mismas, ya que promueven actividades y un valioso uso de recursos para la representación del conocimiento [Kinsella et al. 2009 ; Gemünden, Ritter, & Walter, 1997]. Así, al hablar de una red social, entendemos que es un conjunto de individuos conectados entre sí por algún tipo de relación o parentesco que comparten un tipo de información. La Figura 2:4 representa un esquema conceptual para la representación de un usuario conectado con un conjunto de individuos en una red social, para su representación, el gráfico se ha realizado mediante la aplicación *Cognos.Social*⁸.

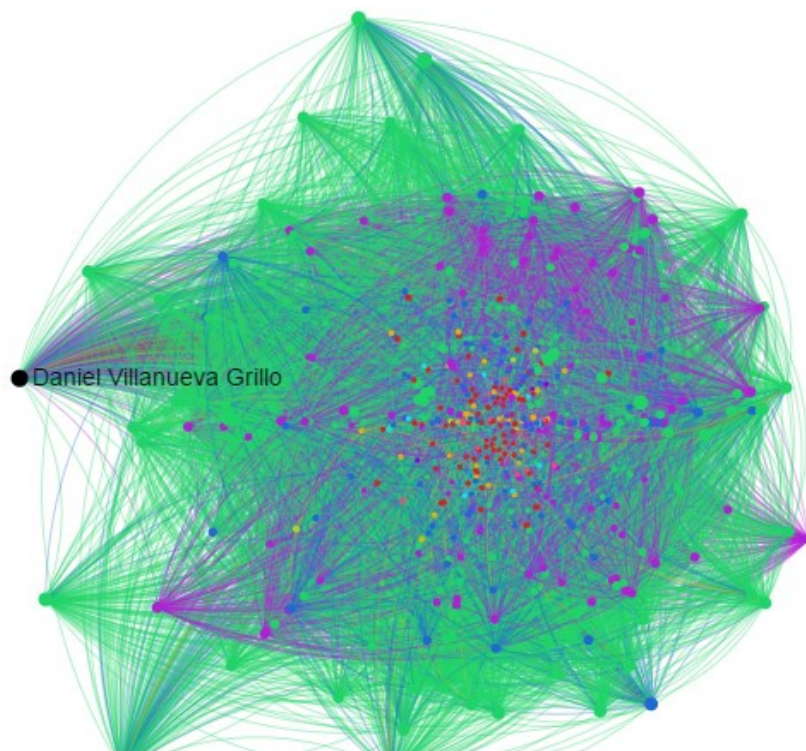


Figura 2:4 Ejemplo de representación esquemática de conjunto de individuos conectados entre sí en una red social.

⁸ <http://social.bal.do/>

El progreso tecnológico y la innovación han alterado la naturaleza de las redes sociales, integrando nuevas formas de comunicación que antes eran imposibles de llevarlas a cabo [Weaver & Morrison, 2008]. Además, las redes sociales difieren con respecto a su tamaño y a la heterogeneidad, un cambio importante en la funcionalidad para representar el conocimiento entre la gente [Mika 2005 ; Garton, 1997]. Por ejemplo, dentro de un contexto comercial, las redes sociales ofrecen beneficios significativos, incluyendo un aumento del conocimiento sobre el valor económico en las organizaciones [Stephen & Toubia, 2010]. Por lo tanto, la representación del conocimiento en las redes sociales son importantes para la supervivencia de las pequeñas empresas y, sugestivas para competir con las grandes empresas [Copp & Ivy, 2001; Pitt & Merwe, 2006].

Una de las redes sociales populares actualmente es Twitter, que es un sitio de microblogging donde los usuarios intercambian mensajes cortos llamados "tweets". Los tweets son mensajes de texto de hasta 140 caracteres que proponen un enfoque intuitivo, de dominio específico, sin supervisión, basada en el léxico que estima el nivel de intensidad emocional contenida en el texto [Paltoglou & Thelwall, 2012].

El sitio cuenta con 500 millones de usuarios registrados que publican 400 millones de tweets por día según el ranking en el mundo Alexa⁹. La plataforma de Twitter para la comunicación rápida se dice que es una plataforma de comunicación vital en los acontecimientos recientes [Morstatter et al. 2013]. Además, Twitter es un sitio de Microblogging que genera un flujo constante de comunicación, algunos de los cuales se refiere a eventos de interés general [Thelwall, 2011].

En este sentido, Twitter es utilizado como una fuente de información para analizar y monitorizar su contenido, además para el desarrollo de distintas herramientas y sistemas que utilizan los contenidos que los usuarios mencionan y opinan a través de la red social. Por lo tanto, los sistemas clasifican y categorizan la información para realizar un análisis sobre los contenidos para distintos dominios. Por comentar alguno, uno de estos sistemas es Twittiment¹⁰, una herramienta de monitorización y análisis sobre los contenidos en

⁹ <http://www.alexa.com/siteinfo/twitter.com>

¹⁰ <http://twittiment.com/>

Twitter. A modo de ejemplo, la Figura 2:5 se muestra una gráfica en respuesta al análisis realizado a través del sistema.

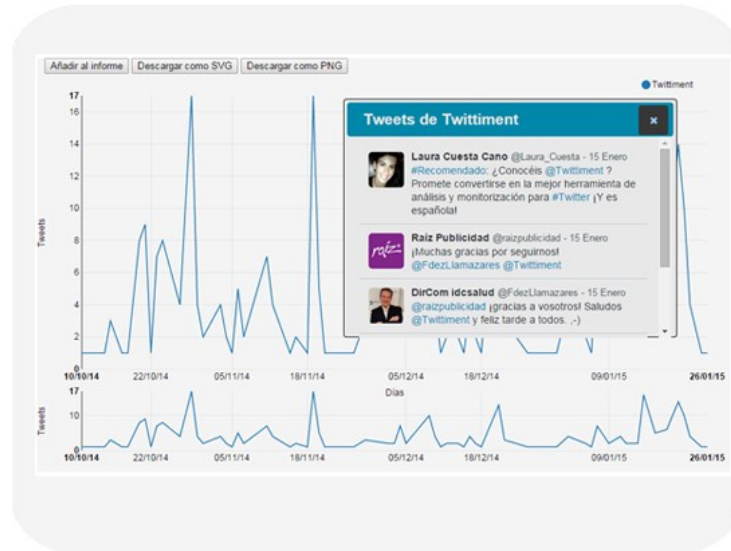


Figura 2:5 Ejemplo de análisis utilizando una herramienta de monitorización en Twitter [e.g., Twittiment].

Por lo tanto, mediante la publicación de un tweet, es posible conocer el comportamiento de la información y, además tiene un componente afectivo, en el sentido de las sentencias o las intenciones de acuerdo con [Tumasjan et al., 2010]. En este sentido, las redes sociales tradicionales han implicado interacciones personales de los seres humanos a través del tiempo, de la difusión de la información y la propagación del conocimiento [Kimball & Rheingold, 2000]. Para ello, existen técnicas y herramientas específicas para la recuperación y representación del conocimiento, tales como, el (PLN) procesamiento del lenguaje natural que realiza una búsqueda con el fin de recuperar un conjunto de información específica para un usuario [Voorhees, 1999] o la recuperación por medio de “Data mining”, que es ampliamente utilizado en diversos dominios para identificar productos y aplicaciones con la finalidad de encontrar patrones o principios entre las grandes cantidades de datos de información [Lehmann et al., 2005]. Cabe señalar, que la

investigación de la presente tesis doctoral se centra en la utilización de ontologías de dominio que ofrecen información estructurada sobre un tema determinado [Monachesi & Markus, 2010]. No obstante, a continuación se incluye un apartado sobre el procesamiento de lenguaje natural que, aunque en esta investigación no se hace uso de esta tecnología, ya que es un proceso basado en un corpus de palabras permite entender el valor de la técnica usada en los medios sociales. Por lo tanto, es importante mencionarla, ya que actualmente la disciplina es utilizada para la recuperación de la información en distintos campos de aplicación, entre ellos, los medios sociales.

2.1.5. Procesamiento de lenguaje natural

El procesamiento de lenguaje natural es una disciplina que consiste en la utilización de un lenguaje natural para comunicarnos con la computadora, debiendo esta entender las oraciones que le sean proporcionadas. El uso de estos lenguajes naturales facilita el desarrollo de programas que realicen tareas relacionadas con el lenguaje o bien, desarrollar modelos que ayuden a comprender los mecanismos humanos relacionados con el lenguaje [Vásquez et al. 2009]. PLN es el método computarizado para el análisis de texto que se basa tanto en un conjunto de teorías y un conjunto de tecnologías. Y, al ser un área muy activa de investigación y desarrollo, no hay una única definición acordada que satisfaga a todos, pero hay algunos aspectos que formarían parte de la definición de cualquier persona con conocimientos, una definición propuesta por Elizabeth D. Lidd es: El PLN es una gama teóricamente de técnicas computacionales para el análisis y la representación de los textos de origen natural en uno o más niveles de análisis lingüístico para el propósito de lograr una similitud a la humana en procesamiento del lenguaje para una variedad de tareas o aplicaciones [Liddy 2001]. Por otro lado, PLN es considerado una disciplina dentro del área de la Inteligencia Artificial y la Lingüística, investiga y formula mecanismos computacionalmente efectivos que faciliten la interrelación hombre/máquina y permitan una comunicación mucho más fluida y menos rígida que los lenguajes formales. [Ovchinnikova 2012].

Las primeras aplicaciones del PLN se dieron durante el período de 1940-1960, teniendo como interés fundamental la traducción automática. Los experimentos en este sector, basados en la substitución de palabra por palabra, obtuvieron resultados rudimentarios. Por lo tanto, surgió la necesidad de resolver ambigüedades sintácticas y semánticas, y asimismo la consideración de información contextual. La carencia de un orden de la estructura oracional en algunas lenguas, y la dificultad para obtener una representación tanto sintáctica como semántica, fueron los problemas más relevantes. Afrontándolos se dio paso a una concepción más realista del lenguaje en la que era necesario contemplar las transformaciones que se producen en la estructura de la frase durante el proceso de traducción.

Uno de los intentos más antiguos programas conocidos en la comprensión del lenguaje natural es el programa “ESTUDIANTE” desarrollado por Bobrow, en su proyecto de tesis. Este fue un programa, que podría resolver los problemas de álgebra de secundaria que plantea en una forma de lenguaje natural [Bobrow 1964]. Actualmente, PLN es utilizado en diferentes campos de investigación relacionados con la recuperación de la información, tales como, robótica [Perzanowski, Schultz, and Adams 1998], medicina [Friedman and Hripcsak 1999], finanzas [Costantino 1997], medios sociales [Verma et al. 2011], ya que proporciona soluciones en la información a través del análisis léxico y sintáctico del texto cuyo objetivo final es facilitar el acceso a la información [Allan 2004 ; Perez-Carballo and Strzalkowski 2000].

El PLN se concibe como el reconocimiento y utilización de la información expresada en lenguaje humano a través del uso de sistemas informáticos. En su estudio intervienen diferentes disciplinas tales como lingüística, ingeniería informática, filosofía, matemáticas y psicología. Debido a las diferentes áreas del conocimiento que participan, la aproximación al lenguaje en esta perspectiva es también estudiada desde la llamada ciencia cognitiva. Tanto desde un enfoque computacional como lingüístico se utilizan técnicas de inteligencia artificial, tales como:

- Modelos de representación del conocimiento y de razonamiento,
- Lenguajes de programación declarativos,
- Algoritmos de búsqueda y estructuras de datos.

En muchas aplicaciones del PLN los objetivos del análisis apuntan hacia el procesamiento del significado. En los últimos años las técnicas de procesamiento semánticos han experimentado avances significativos, resolviendo los diversos problemas fundamentales para distintos dominios. Sin embargo, para esta investigación no se hace uso de esta disciplina sino se centra en la utilización de ontologías de dominio que permiten explotar la representación del conocimiento y la búsqueda de soluciones en la información estructurada sobre un tema determinado. Además, porque las ontologías han llegado a ser comunes en la Web semántica y su éxito se fundamenta en la reutilización eficaz y eficiente del conocimiento para construir representaciones abstractas en el significado de la información.

2.1.6. Ontologías

Las ontologías has llegado a ser comunes en la Web semántica, uno de los factores claves para su éxito es la reutilización eficaz y eficiente del conocimiento [Simperl, 2009]. El término "Ontología" fue introducido a la informática como un medio para formalizar el tipo de cosas de las que se pueden hablar acerca de en un sistema o de un contexto. Con una larga tradición en la filosofía, donde "Ontología" denota, "el estudio del ser o existencia" [Cahn, 2012 ; Advani, Tu, & Musen, 1997 ; Arpírez & Gómez-Pérez, 2000]. No fue sino hasta finales de la década de los 80s que es adoptado por la IA como un término para compartir y reutilizar conocimiento. A mitad de los 90s se incorpora a la ingeniería Web para la inclusión de descripciones semánticas explícitas de recursos (contenidos y servicios), [Berners-Lee et al., 2001]. No obstante, una de las definiciones más extendidas y aceptas especialmente en el campo de la IA sobre el concepto de ontología dentro del contexto de la reutilización del conocimiento, es la de T. R. Gruber, que define como una especificación explícita de una conceptualización [Gruber, 1993].

Una ontología describe diferentes propiedades y relaciones entre las entidades del dominio, según [Hoekstra, 2009]. Además, las ontologías pueden utilizar un proceso de inferencia que permite mejorar los SBC, añadiendo una información relevante [Raimond et al., 2009]. Al mismo tiempo, existen múltiples lenguajes para desarrollar ontologías,



Actualmente existen distintos dominios en la Web semántica que hacen uso de las ontologías para la representación del conocimiento con diversas aplicaciones [Qin & Atluri 2009]. Por ejemplo, en las comunicaciones [e.g., Lemos et al., 2012], el cine [e.g., J. Li, Ding, Shi & Zhang, 2010; Colombo-Mendoza, 2015], plataformas Web [e.g., Guha & McCool, 2003], redes móviles [e.g., Kobeissy, 2009], vigilancia [e.g., Tani & Lablack, 2014], en el dominio financiero [e.g., Valencia-Garcia & Garcia-Sánchez, 2009] y [e.g., Rodríguez-García, 2014], en la ingeniería civil urbana [e.g., Teller & Keita, 2007], en el dominio inversiones [e.g., Gonzalez-Carrasco 2012], industrial [e.g., Garcia-Crespo, 2010], turismo [e.g., García-Crespo et al., 2011], negocios [e.g., Osterwalder & Pigneur, 2004], medicina [e.g., Rodríguez-González et al., 2012], por citar algunos de los dominios en los que se aplican las ontologías.

La ontología formal se especifica mediante una colección de nombres, conceptos y tipos de relaciones organizados en un orden parcial a través de relaciones, las cuales son llamadas clases, estas clases se describen a partir de propiedades que representan las características sobre los atributos y relaciones de las clases, adicionalmente estas características tienen restricciones como tipo, subtipo, etc. [Lemos et al., 2012; López-Cuadrado, 2012]. Finalmente, se tienen instancias (elementos identificables) que construirán los individuos concretos que representa la ontología [Muntjewerff & Bredeweg, 1999].

Para el desarrollo de una ontología se requiere definir clases que forman el dominio, organizar las clases en una jerarquía taxonómica, definir las propiedades de cada clase indicando las restricciones de sus valores y, a su vez, asignando valores a las propiedades para crear instancias, una vez que se tiene una especificación del conocimiento se puede analizar utilizando métodos formales [Noy & Klein, 2004]. Aunque no existe una metodología establecida sobre cómo desarrollar ontologías, se debe tener en cuenta lo siguiente

Una ontología define un vocabulario común para quienes necesitan compartir información en un dominio [Tudorache, Vendetti, & Noy, 2008]. Con lo cual, las ontologías contienen definiciones de conceptos básicos y relaciones que pueden ser interpretadas

por una máquina. Pero con toda esta información surge una pregunta, ¿Por qué alguien desearía desarrollar una ontología? Algunas de las razones son las siguientes:

- Compartir el entendimiento común de la estructura de información entre personas o agentes de software.
- Permitir la reutilización de conocimiento de un dominio.
- Explicitar suposiciones de un dominio.
- Separar el conocimiento del dominio del conocimiento operacional.
- Analizar el conocimiento de un dominio.

Por lo tanto, una ontología es una jerarquía de conceptos con atributos y relaciones, que define una terminología consensuada para definir redes semánticas de unidades de información interrelacionadas. Una ontología proporciona un vocabulario de clases y relaciones para describir un dominio, poniendo el acento en la compartición del conocimiento y el consenso en la representación de éste. Por ejemplo, una ontología sobre arte podría incluir clases como pintor, cuadro, estilo o museo y, relaciones como autor de un cuadro, pintores pertenecientes a un estilo artístico u obras localizadas en un museo. Además de las clases y relaciones las ontologías utilizan componentes que sirven para representar el conocimiento de algún dominio [Horrocks, Patel-Schneider, & Harmelen, 2002]. Los componentes se describen a continuación:

- **Conceptos** – Son las ideas básicas que se intentan formalizar. Los conceptos pueden ser clases de objetos, métodos, planes, estrategias, procesos de razonamiento, etc.
- **Relaciones** – Representan la interacción y enlace entre los conceptos del dominio. Suelen formar la taxonomía del dominio. Por ejemplo: subclase-de, parte-de, parte-exhaustiva-de, contiene-un, conectado-a, etc.
- **Funciones** – Son un tipo concreto de relación donde se identifica un elemento mediante el cálculo de una función que considera varios

elementos de la ontología. Por ejemplo, pueden aparecer funciones como categorizar-clase, asignar fecha, etc.

- **Instancias** – Se utilizan para representar objetos determinados de un concepto.
- **Axiomas** – Son teoremas que se declaran sobre relaciones que deben cumplir los elementos de la ontología. Por ejemplo: “Si A y B son de la clase C, entonces A no es subclase de B”, “Para todo A que cumpla la condición C, A es B”, etc.

Teniendo en cuenta los componentes que integran una ontología, es posible crear una base de conocimientos definiendo las instancias individuales de esas clases, precisando los valores específicos de los slots y restricciones adicionales sobre los mismos [Lame, 2005]. En síntesis, una ontología es un sistema de representación del conocimiento que resulta de seleccionar un dominio o ámbito del conocimiento, y aplicar sobre él un método con el fin de obtener una representación formal de los conceptos que contiene y de las relaciones que existen entre dichos conceptos [Zavitsanos & Paliouras, 2007]. Además, una ontología se construye en relación a un contexto de utilización [Klein & Fensel, 2001]. Con lo cual, quiere decir que una ontología especifica una conceptualización o una forma de ver el mundo, por lo que cada ontología incorpora un punto de vista. Además, una ontología contiene definiciones que nos proveen del vocabulario para referirse a un dominio [Kobeissy, 2009]. Estas definiciones dependen de la utilización de un lenguaje para describir todas las conceptualizaciones (definiciones, categorizaciones, jerarquías, propiedades, herencia, etc.) de una ontología, que a su vez, pueden ser procesables por máquinas [Schlenoff & Prestes, 2012; Nilsson & Muradore, 2009]. Por lo tanto, a diferencia de la disciplina del procesamiento de lenguaje natural, en la presente investigación se utilizan las ontologías, ya que permiten construir nuevos modelos conceptuales para representaciones en el significado de los contenidos, a su vez, permiten explotar la semántica de la información disponible en la Web desde distintos paradigmas.

2.2. Clasificación de modelos para Sistemas de Recomendación

2.2.1. Modelos de recomendación

Los Sistemas de Recomendación tienen la capacidad de desarrollar un razonamiento conocido actualmente por sus aplicaciones en diversos dominios [Drachsler, Hummel, & Koper, 2007]. En estos sistemas se argumenta que la extracción de filtrado de información más significativa, son los datos de los modelos conceptuales computacionales en las que están basadas sus recomendaciones. Este es un reto que debe abordarse para que los SR sean desarrollados de manera más comprensible, eficaz y aceptable [Mcsherry, 2005].

Para los distintos modelos conceptuales existen diferentes enfoques y técnicas clasificadas que permite identificar su uso y aplicación sobre un distintos dominios [Rodríguez-Gonzalez et al., 2012]. Por ejemplo, el área específica de modelos de dominios cruzados CD (Cross-Domain) es, en estos momentos uno de los temas de investigación emergente en el área de los SR, acordado por [Lops Pasquale, 2011], otro tema son los modelos basado en conocimiento con técnicas inteligentes para los SR [Rodríguez-González, 2012].

Estos sistemas utilizan, además de las valoraciones de los usuarios sobre los productos, información descriptiva de cada producto. El contenido almacenado puede ser de distinta naturaleza. Por ejemplo, una descripción mediante lenguaje natural [e.g., Degemmis, Lops, & Semeraro, 2007 ; e.g., Resnick, Iacovou, & Suchak, 1994 ; e.g., Reilly et al., 2004] o, mediante atributos de los productos con características y valores asociados a los mismos [e.g., Bezerra & de A.T. de Carvalho, 2004 ; e.g., Leite Dantas Bezerra & Tenorio de Carvalho, 2010]. Por otro lado, CD ha utilizado el término “dominio” en algún momento para hacer referencia a los tipos de elementos. Por ejemplo, películas vs libros

o, en otros casos, a los grupos de elementos similares con características comunes, películas vs programas de televisión [Milicevic, Nanopoulos, & Ivanovic, 2010].

En la literatura existen propuestas a las soluciones de problemas clásicos sobre los distintos modelos utilizados en los SR, la literatura los clasifica básicamente en los siguientes tipos según [Adomavicius & Tuzhilin, 2005 ; Lu et al. 2015].

- Modelos para Sistemas de Recomendación Colaborativos (Collaborative Recommender System, SRC), [e.g., Goldberg et al., 1992].
- Modelos para Sistemas de Recomendación Demográficos (Recommender System Demography, SRD), [e.g., Christensen & Schiaffino 2011].
- Modelos para Sistemas de Recomendación Basados en Contenido (System Recommender Content Based, SRBC), [e.g., Rana & Jain 2012].
- Modelos para Sistemas de Recomendación Basados en Conocimiento (System Recommender Knowledge Based, SRKB), [e.g., Martínez et al., 2008].
- Modelos para Sistemas de Recomendación Híbridos (System Recommender Hybrid, SRH), [e.g., Bao, Bergman, & Thompson, 2009].
- Modelos para Sistemas de Recomendación Basado en Factores Económicos (System Recommender Economic Factor Based, SREFB), [e.g., Lee, 2004].

Los SRC aparecieron a mediados de los años 90 acordado por [Goldberg et al., 1992] y su funcionamiento establece modelos mediante la información del resto de los usuarios que hay en el sistema. Por ejemplo, Resnick et al., presentan un sistema que

ayuda a las personas a tomar decisiones sobre la base de opiniones de otras personas [Resnick et al., 1994]. Dichos sistemas se basan en la idea de usuarios con características dadas que tengan similitud a otros usuarios con características demográficas similares. Además, estos modelos se aplican a dominios concretos, acordado por [Rich, 1983; Schiaffino & Amandi, 2009; Vozalis & Margaritis, 2007].

Los modelos para SRBC contienen información sobre las características de cada producto e intentan extraer las relaciones entre estas y las preferencias de los actores mediante un perfil a partir de productos valorados [Rana & Jain, 2012], estableciendo una predicción mediante distintas técnicas como, redes bayesianas [e.g., Simonsson, Lagerström, & Johnson, 2008] y redes neuronales [e.g., Christakou, Vrettos & Stafylopatis, 2007]; dichas técnicas son utilizadas mayormente para procesar la idoneidad de un producto preferente para los actores.

Los modelos para SRKB, utilizan una base de conocimiento para describir los distintos productos que satisfacen las necesidades de los actores [Martínez et al., 2008], de manera que, el sistema encuentra el producto que más se ajusta a las necesidades que los actores han especificado [Bridge & Ferguson, 2002].

Asimismo, la información adquirida sobre las características de los productos resulta importante, ya que permite obtener una información detallada sobre las relaciones de los usuarios con distintos productos, con lo cual, la caracterización permite conocer necesidades específicas; y a su vez, resulta difícil, trasladar un modelo de un dominio a otro, ya que la mayoría de los modelos descritos son desarrollados para trabajar en dominios específicos (i.e., películas, libros, música, noticias, automóviles). Por lo tanto, los procesos de la ingeniería del conocimiento, utiliza dichos modelos para lograr resultados favorables, modelos conceptuales con un mayor alcance en el campo de aplicación. Para ello, la investigación de la presente tesis doctoral se centra en el desarrollo de un modelo conceptual que pueda ser aplicado a múltiples dominios sin cambiar su base de conocimiento original.

Los SRH emplean modelos que son una mejora para las recomendaciones, acordado por [Barragáns-Martínez et al., 2010]. Además, mediante una hibridación de distintas técnicas [Robin Burke, 2002], se realiza una combinación de A y B para obtener

un resultado que combina el comportamiento de ambos y, genera de esta manera una respuesta para el dominio o entorno específico [Bao et al., 2009].

Los SREFB basan sus modelos en la información que definen los actores, valorando la utilidad de los productos. Estos sistemas evalúan los productos y recomiendan un conjunto de productos con mayor valor de utilidad por los actores [Lee, 2004].

Esta clasificación de modelos y técnicas son una tendencia actual en el desarrollo de los sistemas comentados con anterioridad, ya que combinan modelos con características de recomendaciones basadas en el contenido estructurado y recomendaciones de filtrado de información. Sin embargo, las técnicas mencionadas en la clasificación generan modelos específicos para un solo dominio, de manera que delimitan al modelo conceptual para ejecutarse en distintos dominios. Por ejemplo, una persona que quisiera escoger entre los nuevos lanzamientos de música, obras de teatro, películas, libros o programas de televisión, como resultado al uso de los modelos descritos podría obtener solo unas determinadas opciones. Por lo tanto, es importante el diseño de nuevos modelos para distintos dominios en los medios sociales.

2.2.2. Modelos multi-dominios existentes en los medios sociales

El mundo moderno de la tecnología de medios digitales y el contenido de datos en la Web y las TS, hace que compartir, colaborar con los demás y conectar entre sí para crear una comunidad más rápida, sea mucho más fácil y más accesible a una población amplia y cada vez más exigente en la información a sus necesidades según [Zhou et al., 2011]. La Web contiene una gran colección de datos sobre sitios tales como, webs sociales, bibliotecas digitales y sitios web conocidos como Wikis, de manera que muchos de los comportamientos sociales pueden reflejarse mediante la integración de estos datos procedentes de medios heterogéneos y, al mismo tiempo, son utilizados cada vez con mayor asiduidad en los SBC, una información detallada que sirva para recomendaciones transparentes a diferentes usuarios [Tiroshi et al., 2012].

Actualmente es bien reconocido que el contenido generado por el usuario (i.e., análisis de productos, etiquetas, foros de debate, entretenimiento y blogs), contiene valiosas opiniones de los usuarios que pueden ser explotadas por distintas herramientas y TS, tales como, RSS (*Really Simple Syndication*), un formato para syndicar o compartir contenido en la Web.

Los medios sociales y los SBC personalizados, pueden beneficiarse mutuamente el uno del otro: por un lado, los medios sociales introducen nuevos tipos de datos públicos, tales como los metadatos: etiquetas, valoraciones, comentarios y relaciones explícitas de los actores, de manera que dicho contenido puede ser utilizada por SBC para mejorar las recomendaciones según [Xiang & Yang, 2009a]. Por otra parte, los modelos usados en las TS de dichos sistemas, pueden desempeñar un papel clave en el éxito de las aplicaciones de los medios sociales y la Web social en su conjunto, lo que garantiza que a cada usuario se le presente el contenido mucho más atractivo y relevante adaptado a sus necesidades de forma personal y específica [Guy et al., 2010].

Por lo tanto, la personalización y caracterización de los productos es cada vez mayor entre los actores [Franke, Schreier, & Kaiser, 2010]. A un usuario pueden interesarle ciertas propiedades sobre un producto, tales como, el diseño, marca, precio, tecnología, etc. De manera que, la personalización y caracterización de los productos ejerce un sentimiento sobre los actores a la hora de tomar ciertas decisiones [B. Liu, 2011].

Los SBC, ayudan a los actores a elegir sobre un conjunto de características a través de procesos de personalización que satisfagan sus necesidades mediante el acceso a una Web y las TS, lo cual, permite configurar las características de los distintos productos mediante técnicas de personalización de contenidos en función a sus necesidades [Glorot, Bordes, & Bengio, 2011]. A raíz de la proliferación de los medios sociales en la Web, han surgido distintas recomendaciones sobre sitios aplicados a diferentes dominios. Por ejemplo, dentro de las organizaciones y empresas han ganado una popularidad semejante a sus equivalentes en la Web.

La recomendación basada en contenido consiste en recomendar aquellas promociones similares a las más valoradas por el usuario. La forma de estimar cuánto le

gusta una promoción a un usuario consiste en la puntuación de la promoción para ese usuario,

Existen modelos que utilizan técnicas basadas en contenido multi-domino para representar productos en los medios sociales, por consiguiente son usados por SBC que tratan de explotar un conocimiento, presidir una valoración dada por el usuario y, modelar un determinado comportamiento de los usuarios sobre distintos productos [Milicevic et al., 2010]. Por ejemplo, Guy et al., presentan el estudio de un modelo para un SR personalizado en los medios de comunicación, específicamente aplicado a las medios sociales basado en personas y etiquetas, utiliza una técnica de filtrado híbrido [Guy et al., 2009]. Por otra parte, Ha I et al., describen un modelo usado en el dominio del entretenimiento en la Web, el modelo usa técnicas de filtrado colaborativo, usuarios similares (vecinos) que tienen una relación social con los medios de comunicación social; el modelo evalúa la precisión de datos preferentes sobre películas a los distintos actores Ha I et al., [2012]. Por otro lado, Colombo et al., presentan un modelo sensible al contexto basado en las tecnologías de la Web semántica para un sistema de recomendación móvil [Colombo-Mendoza, 2015].

Los modelos aplicados a los SR en los medio sociales juegan un papel importante en la Web semántica, como pueden comprobarse en sitios de Internet de alta calidad, tales como, Amazon¹¹, Netflix¹², iTunes¹³, TripAdvisor¹⁴, Lasf.fm¹⁵ e IMDb¹⁶. Por otra parte, muchas compañías de medios sociales están desarrollando y desplegando SR como parte de los servicios que prestan a sus abonados [Mak, Koprinska, & Poon, 2003].

Liang Hu et al., presentan un modelo sobre la relación de los usuarios en los medios sociales, el modelo proporciona una mayor eficiencia en comparación con los modelos tradiciones [Liang Hu et al., 2013]. Además, Cremonesi et al., presentan algoritmos basados en comunidades para múltiples dominios, evaluando un conjunto de datos creados artificialmente [Cremonesi et al., 2011]. Cabe señalar que ninguno de los

¹¹ <http://www.amazon.es/>

¹² <https://www.netflix.com/global>

¹³ <https://www.apple.com/es/itunes/>

¹⁴ <http://www.tripadvisor.es/>

¹⁵ <http://www.lastfm.es/>

¹⁶ <http://www.imdb.com/>

modelos antes mencionados pueden ser usados en un dominio distinto al que fueron propuestos originalmente, es decir, que los modelos se limitan a trabajar a un solo dominio. Por ejemplo, Netflix sugiere películas y series de televisión, y Last.fm genera recomendaciones personalizadas de artistas musicales y composiciones; en ambos casos, los modelos utilizan la información almacenada para generar recomendaciones que se calculan utilizando calificadores acerca de los elementos, pero solo y exclusivamente en el dominio de destino.

Por lo tanto, no es posible trasladar ambos modelos a un dominio diferente sin cambiar su base de conocimiento, de forma que estos modelos son aplicados a un solo dominio utilizando técnicas basadas en filtrado de información y basados en el conocimiento [Fernández-Tobías et al., 2011].

2.2.3. Modelos formales semánticos multi-dominio

Sobre la base de la clasificación propuesta anteriormente, los modelos semánticos utilizan técnicas Basadas en Contenido (*Content-Based*, CB), Filtrado Colaborativo (*Filtering-Collaborative*, CF) y Basadas en el Conocimiento (*Knowledge-Based*, KB), acordado por [Wang & Kong, 2007]. De esta manera, los modelos semánticos conceptuales aplicados a distintos dominios, hacen uso de estas técnicas para el desarrollo de Sistemas Basados en Conocimiento con enfoques distintos que suponen la integración y gestión del conocimiento mediante el razonamiento basado en casos, una técnica de resolución de problemas que utiliza el conocimiento [Kolodner, 1992], o conocimiento basado en ontologías [Guy et al., 2009]. Por ejemplo, Alejandro et al., presentan una mejora sobre la eficiencia del tiempo en un sistema basado en el razonamiento sobre ontologías [Alejandro Rodríguez et al., 2012]. Maedche, Motik, and Stojanovic presentan un marco de trabajo para la gestión múltiple en la Web semántica que se basa en el modelo de representación de ontologías [Maedche, Motik, & Stojanovic, 2003b]. Guo and J presentan una descripción sobre un enfoque basado en una gran ontología OWL para especificar tipos de razonamiento [Guo & Heflin, 2006]. Además, Bock hace la observación de que los razonadores disponibles no explotan todas las

ventajas de las técnicas de computación, por lo tanto, propone una aplicación en paralelo al razonamiento con ontologías [Bock, 2008]. Por lo tanto, los modelos semánticos son un área de explotación en el campo de los Sistemas de Recomendación.

La Figura 2:7, muestra una clasificación genérica de los modelos semánticos en los medios sociales. Basado en la descripción anterior, estos modelos se clasifican básicamente en seis tipos de técnicas utilizadas en los Sistemas de Recomendación.

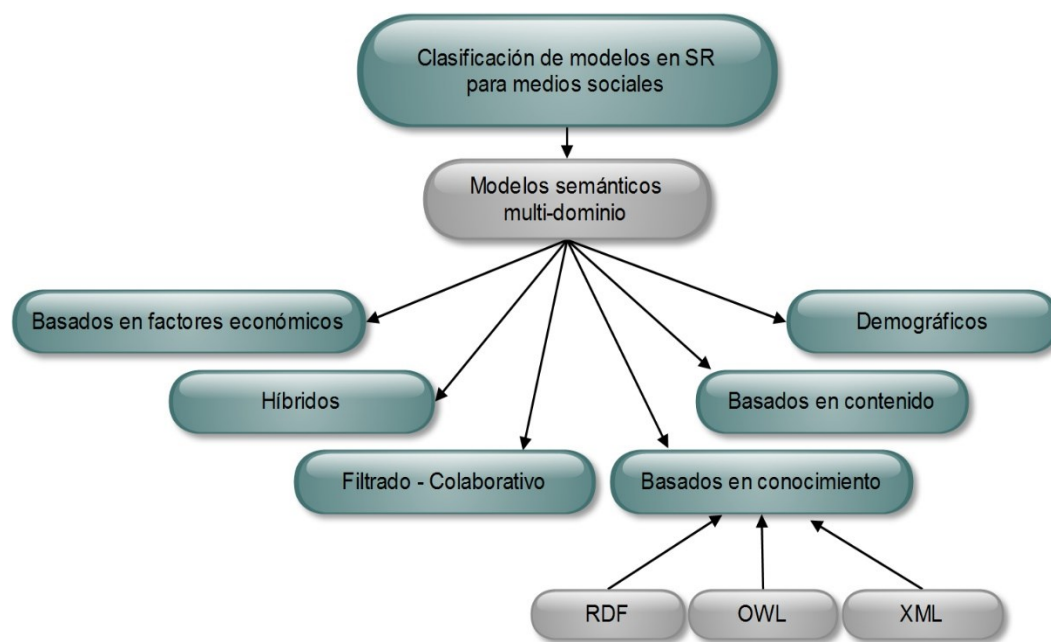


Figura 2:7 Clasificación de técnicas usadas para modelos formales en los medios sociales.

Por otro lado, una forma de especificar el conocimiento sobre un dominio concreto son las ontologías [A Rodríguez-González, 2012b]. Tal y como se ha comentado en la Sección 2.1.5. Por lo tanto, las ontologías pueden ser utilizadas para compartir información entre distintos sistemas, de manera que se estimula la cooperación del contenido, lo cual ha supuesto para esta investigación, trabajar con el surgimiento, desarrollo y concepción de nuevos modelos ontológicos semánticos basados en conocimientos para distintos dominios.

Lu et al., presentan un interesante y exhaustivo análisis en el desarrollo de aplicaciones hasta al día de hoy sobre los SR, agrupando sus aplicaciones en ocho dominios principales. En consecuencia, Lu et al., resume las técnicas de recomendación que se utilizan en relación a los dominios que presentan [Lu et al., 2015]. En el análisis examinan sistemáticamente los SR reportados a través de cuatro dimensiones. Además, el estudio que presentan proporciona una encuesta que apoya a la investigación de la presente tesis doctoral en el sentido del exhaustivo análisis de la evolución de las técnicas y aplicaciones usadas en los SR. En la Tabla 2.1 se muestran el resumen que presenta Lu et al., que además, detallan el análisis de los distintos modelos en Sistemas de Recomendación y las técnicas aplicadas para cada uno de los diferentes dominios.

Tabla 2.1:

Encuesta de estudio para distintos modelos en Sistemas de Recomendación, dominio de aplicación y tipo de usuarios. Tabla propuesta por [Lu et al., 2015].

	Application domain	Technique	Application platform	User type	Period	Reference
A multi-agent e-government system	e-government	KB	Web-based	Individual	2005	[76]
eElections RS	e-government	Fuzzy clustering	Web-based	Individual	2010	[75]
TPLUFIB-WEB	e-government	Fuzzy linguistic modeling, Hybrid, CB, CF	Web-based	Individual	2014	[77]
Smart trade exhibition finder	e-government	CF, Hybrid	Web-based	Business	2007	[74]
A trade exhibition recommender system for e-government	e-government	CF, Hybrid, Fuzzy logic	Web-based	Business	2007	[49, 78]
BizSeeker	e-government	CF, Hybrid	Web-based	Business	2010	[79]
Smart BizSeeker	e-government	CF, Hybrid, Fuzzy sets	Web-based	Business	2013	[10, 80]

An ontology-based product RS	e-business	KB, Bayesian belief network	Web-based	Business	2006	[82]
Auction seller recommender system	e-business	Social network analysis	Web-based	Business	2008	[83]
A negotiation-style recommender	e-business	CI, a simulated-annealing inspired algorithm, greedy algorithm	Web-based	Business	2011	[84]
PB-ADVISOR	e-business	Fuzzy logic, KB	Web-based	Business	2012	[85]
Telecom recommender system	e-business	CF, KB, Hybrid, Fuzzy sets	Web-based	Business	2013	[9]
MusicBox	e-commerce	Social tag, CF, Hybrid	Web-based	Individual	2010	[86]
Wasabi personal shopper	e-commerce	CB, KB	Web-based	Individual	1999	[88]
Consumer electronic products RS	e-commerce	KB, CI, Fuzzy techniques	Web-based	Individual	2007	[48]
A book RS	e-commerce	CB	Web-based	Individual	2000	[89]
MRH	e-commerce	Social network, CI	Web-based	Individual	2011	[90]
A conversational RS	e-commerce	KB	Web-based	Individual	2004	[91]
GRAB	e-commerce	KB, CI	Web-based	Individual	2006	[92]
A supermarket product RS	e-commerce	CB, CI	Mobile-based	Individual	2001	[93]
Fab	e-library	CB, CF, Hybrid	Web-based	Individual	1997	[96]

CYCLADES	e-library	CB, CF, Hybrid	Web-based	Individual, Group	2005	[97]
University digital library Recommender system	e-library	Hybrid, Fuzzy linguistic modeling	Web-based	Individual	2009-2011	[50, 95, 98, 99]
An e-learning recommender agent	e-learning	KB, Rule mining	Web-based	Individual	2002	[100]
PLRS	e-learning	CB, KB	Web-based	Individual	2004	[101]
AHA!	e-learning	Web usage mining	Web-based	Individual	2009	[102]
FIRT	e-learning	KB, CI, Fuzzy item response theory	Web-based	Individual	2004	[103]
FIRT	e-learning	KB, CI, Fuzzy item response theory	Web-based	Individual	2008	[104]
IWT	e-learning	KB	Web-based	Individual	2014	[105]
CourseAgent	e-learning	KB	Web-based	Individual	2006	[106]
RSPP	e-learning	Ontology	Web-based	Individual	2013	[108]
Willow system	e-learning	KB	Web-based	Individual	2014	[109]
PSDLO	e-learning	KB, CB	Web-based	Individual	2009	[110]
Entrée	e-tourism	KB	Web-based	Individual	1996	[111]
EntreeC	e-tourism	KB, CF	Web-based	Individual	2002	[14]
Restaurant directory services agent	e-tourism	Context-aware	Mobile-based	Individual	2004	[112]
CATIS	e-tourism	Context-aware	Mobile-based	Individual	2003	[113]
REJA	e-tourism	CF, KB	Web-based	Individual	2009	[114]

PSiS	e-tourism	CF CB clustering associative	Web-based	Individual	2013	[115]
SigTur/E-Destination	e-tourism	CF CB demographic context	Web-based	Individual	2013	[116]
SMARTMUSEUM	e-tourism	Ontology context information filtering	Mobile-based	Individual	2013	[117]
iTravel	e-tourism	CF context	Mobile-based	Individual	2013	[118]
DIETORECS	e-tourism	KB	Web-based	Individual	2003	[119]
Moleskiing	e-tourism	Trust	Web-based	Individual	2005	[120]
MASTROCARONTE	e-tourism	KB, Context-aware	Mobile-based	individual	2003	[121]
SPETA system	e-tourism	Social networks, Semantic Web, Context-aware	Mobile-based	Individual	2009	[122]
Traveller	e-tourism	CB, CF, Demographic information	Web-based	Individual	2009	[123]
Tag recommender system	e-resource	CF	Web-based	Individual	2011	[124]
FolkRank	e-resource	CF, CB	Web-based	Individual	2006-2007	[125, 126]
Tag recommender system	e-resource	CF	Web-based	Individual	2009	[127]
PTV	e-resource	CB, CF	Web-based	Individual	2000	[128]
TiVo	e-resource	Clustering, CF	TV-based	Individual	2004	[129]

TV recommender system	e-resource	Information retrieval clustering	TV-based	Individual	2010	[130]
MBCF-based program RS	e-resource	CF	TV-based	Individual	2011	[131]
queveo.tv	e-resource	CB, CF	TV-based	Individual	2010	[132]
TV program RS	e-resource	Bayesian classifier, Decision tree	TV-based	Individual	2004	[134]
AVATAR	e-resource	Semantic analysis, CB, CF	TV-based	Individual	2006	[135]
AMALTHAEA	e-resource	Information filtering, Information retrieval	Web-based	Individual	1997-1998	[136, 137]
ifWeb	e-resource	CB	Web-based	Individual	1997	[138]
News Dude	e-resource	IR, Bayesian classifier	Web-based	Individual	1999	[30]
Eigentaste	e-resource	CF, Dimensionality reduction	Web-based	Individual	2001	[139]
GroupLens	e-resource	CF	Web-based	Individual	1994-1997	[21, 140]
Foxtrot	e-resource	k-Nearest classification	Web-based	Individual	2002	[141]
WinPUM	e-resource	Graph based clustering	Web-based	Individual	2010	[142]
Web-page recommender	e-resource	Ontology KB	Web-based	Individual	2013	[143]
Lifestyle Finder	e-resource	Demographic information	Web-based	Individual	1997	[144]

ACR News	e-resource	CB, Clustering	Web-based	Individual	2000	[133]
ArgueNet	e-resource	CB	Web-based	Individual	2004	[145]
PocketLens	e-resource	CF, model-based	Web-based	Individual	2004	[146]
CinemaScreen	e-resource	CF, CB	Web-based	Individual	2006	[147]
Flycasting	e-resource	CF	Mobile-based	Individual	2001	[148]
Smart Radio	e-resource	CF	Radio-based	Individual	2001	[149]
RACOFI	e-resource	Semantic Web, CF	Web-based	Individual	2003	[150]
Foafing the Music	e-resource	Social network, CB	Mobile-based	Individual	2005	[151]
CBCF	e-resource	CB, CF	Web-based	Individual	2002	[152]
CoFoSIM	e-resource	Multi-criteria decision-making, CF	Mobile-based	Individual	2010	[153]
I-SPY	e-group		Web-based	Group	2003-2006	[154-156]
GRec_OC	e-group	CB, CF	Web-based	Group	2010	[157]
CAPS	e-group	CB	Web-based	Group	2003	[158]
MusicFX	e-group	CB	Radio-based	Group	1998	[159]
Flytrap	e-group	CB	Radio-based	group	2002	[160]
Adaptive radio	e-group	CF	Radio-based	group	2005	[161]
PolyLens	e-group	CF	Web-based	group	2002	[69]
Glue	e-group	CF, TKI	Web-based	group	2009	[162]

Pocket restaurant finder	e-group	CB	Web-based	group	2002	[164]
CATS	e-group	CB, Critiquing synchronous	Web-based	group	2006	[73, 165]
INTRIGUE	e-group	Weighted average	Web-based	group	2003-2005	[163, 166]
PETs	e-group	Demographic-based CB, CF	Web-based	group	2009	[167]
e-Tourism	e-group	Demographic, CB	Web-based	group	2011	[168]
DCOP-based multiagent	e-group	Agent	Web-based	group	2008	[169]
TDF	e-group	Asynchronous	Web-based	group	2004	[72, 170]
FIT	e-group	CB	TV-based	group	2004	[171]
TV4M	e-group	CB	TV-based	group	2006	[172]
TV programme recommender	e-group	Classifier	TV-based	group	2009	[173]

En consecuencia y, a partir del análisis realizado en la Tabla 2.2. Lu et al., presentan una segunda tabla que muestra un resumen sobre las técnicas aplicadas para cada dominio, además comentan su importancia de aplicación en los SR.

Tabla 2:2:

Resumen de la encuesta para Sistemas de Recomendación, técnicas y tipo de usuario. Tabla propuesta por [Lu et al., 2015].

Techniques Domains	CB	CF	KB	Hybrid	Computational Intelligence	Social Network	Context Aware	Group Aggregation	No. of listed references
E-government	1	5	1	5	4				9
E-business		1	3	3	4	1			5
E-commerce/E-shopping	3	1	4	1	4	2			8

E-library	2	2		3	1				6
E-learning	2		11		2				10
E-tourism	5	9	9	9	3	2	11		18
E-resource	9	16	6	15	8	1	1		27
E-group activity	9	5	2	5	1			2	21
Total	31	39	36	41	27	6	12	2	104

En el análisis que se realiza, se reafirma que a pesar de que las aplicaciones y técnicas empleadas en cada uno de los Sistemas de Recomendación y, el desarrollo que han obtenido, todavía hay algunos temas que requieren mayor investigación, como lo es, los modelos semánticos multi-dominio utilizados en SR aplicados a los medios sociales, ya que con el surgimiento de las nuevas aplicaciones se ofrecen distintos servicios a partir de la información almacenada en la Web. Por otro lado, se concluye sobre la evolución de los SR aplicados a distintos dominios, la investigación básicamente se distingue por el desarrollo de aplicaciones en el mundo real, por un lado la examinación de modelos y técnicas más destacables y utilizadas en los SR, y por otro lado, la examinación sobre los distintos dominios de aplicación, dejando a un lado, los modelos semánticos, por lo que permite la oportunidad de profundizar en las recomendaciones mediante el filtrado de información semántica en SR en los medios sociales.

En consecuencia, el análisis detallado que presenta Lu et al., permite identificar y conocer los distintos modelos utilizados actualmente en los SR. No obstante, cabe mencionar que la encuesta que presenta no refuerza el uso de los modelos semánticos utilizados en los Sistemas de Recomendación para medios sociales. Sin embargo, su encuesta es un soporte para el análisis de la presente investigación, ya que permite conocer la originalidad y los recursos existentes en el área de los SR. De hecho, la presente investigación se centra en la importancia de modelos semánticos utilizados en los SR aplicados en el área de los medios sociales.

Por lo tanto, es necesario realizar una comparación de los principales modelos semánticos utilizados en los SR para medios sociales. La revisión de los modelos es

basada en veinte artículos de revistas y conferencias importantes con características relevantes. Por ejemplo, la mayor parte de los documentos antes mencionados enfatizan a una clasificación usada en múltiples dominios de modelos semánticos para Sistemas de Recomendación en los medios sociales. Además, al hacer una revisión de la literatura sobre estos temas, la Tabla 2.3 muestra una lista de autores con características, tales como, que describen modelos semánticos aplicados a diferentes dominios. Asimismo, utilizan las técnicas de CF, KB y CB. Por otra parte, también se describe el dominio donde se aplican estos modelos así como sus características. Igualmente, se especifica si el modelo es multi-dominio y/o semántico. Finalmente se detalla una descripción breve del método usado por los autores para evaluar los distintos modelos.

Tabla 2:3:

Encuesta de estudio para distintos modelos formales semánticos en Sistemas de Recomendación.

Authors	Domain and Characteristics	Multi-Domain Models Yes / No	Semantic Models Yes / NO	Evaluation
Agius & Angelides (1999)	Multimedia Systems, Video and Audio m-frames, (KB) and (CB)	Yes	Yes Semantic m-frames	The model serves for the development of intelligent multimedia tutoring systems
Liang Hu et al. (2013)	Social networking sites model over the triadic relation user-item-domain, (KB) and (CF)	Yes	No	The model provides greater efficiency compared to traditional models
Glass, Marx, Schmidt & Sivrikaya (2010)	Multimedia content, IPTV Systems High level, Linked open data (CB) and (KB)	Yes	Yes Semantic Ontology	Semantic data integration IPTV application
Lops Pasquale (2011)	Knowledge Sources Techniques for representing structured data, Vector Space Model (VSM), (CB)	No	No	Models review Information Retrieval research area

Di Noia et al. (2012)	Web applications, Films movie domains, Linked Open Data, Using SPARQL (CB) and (KB)	No	Yes Semantic Web	Precision and recall metrics. The model leverages the knowledge encoded in semantic datasets
Huming & Weili (2010)	Network technology Tourism Rankboost Algorithm (CF)	No	No	Algorithm can make it convenience for the consumer to book desired hotel
Raimond et al. (2009)	TV programme RDF Services (CB) and (KB)	No	Yes Semantic Ontology	The model is used to a specific framework
Ignacio Fernández et al. (2011)	Artificial Intelligence framework for (CD) Linked Data (KB)	Yes	Yes Semantic Network	Discover semantic paths between places of interest and music artists.
Cremonesi, Tripodi & Turrin (2011)	Multiple domain algorithms based on Neighborhood. (CD)	Yes	No	New algorithms, evaluated on a multi-domain dataset artificially created.
Ruiz-Montiel & Aldana-Montes (2009)	Context Ontology User modelling, (CB) and (KB)	No	Yes Semantic Ontology	Precision recall, The model improves the performance of recommendation
Shapira, Rokach & Freilikhman (2013)	Social Networks, Facebook, movies, TV shows and music Items Data sets, (KB)	No	No	Compared the results with two baseline state of the art collaborative filtering algorithms
Tsatsou, Mezaris & Kompatsiaris (2012)	Networked media, Digital media, Linked to the Web, open data, (CB) and (KB)	No	Yes Semantic Ontology	The model selects an efficient knowledge base on the net

Blanco-Fernández et al (2011)	Online shopping Filtering approaches framework (CB) and (KB)	No	Yes Semantic Ontology	Exploits the semantics formalized in an ontology to time functions
Ghani & Fano (2002)	Online Shopping, apparels in Web sites Learning algorithm, (CB) and (KB)	No	Yes Knowledge Based	Build a model of the user from an original model
Belk, Germanakos & Tsianos (2010)	Web Applications ecommerce, Web personalization, (CB) and (KB)	Yes	Yes Ontological	Web Personalization based on an ontological model
Hofmann (2004)	Model-based algorithms (CF) (KB)	No	Yes Semantic analysis	The model discover user communities and interest profiles
Mobasher, Jin & Zhou (2004)	Web Algorithm (CF) and (KB)	Yes	Yes Semantic knowledge ontology	The algorithm create a combined similarity measure and generate predictions
Boim, Milo & Novgorodov (2011)	Movie recommendation Framework DiRec is implemented in Java and PHP, CF-based, (KB)	No	Yes Semantic information on items	Demo plug-in provides a solution in common scenarios where semantic information is unavailable
Xiang & Yang (2009)	Multimedia Analyze how temporal features influence CF Recommender systems (CF)	No	No	Experimental analysis, the model show the prediction accuracy. Model of time Response

Bedi, Kaur & Marwaha (2007)	Semantic Web Tourism Ontology (CB) and (KB)	Yes	Yes Knowledge based	The model work seamlessly across different versions of ontologies
-----------------------------	--	-----	---------------------------	--

En su obra [Agius & Angelides, 1999], presentan un modelo semántico al que llama multi-dominio aunque realmente su aplicación es para un solo dominio, el de las comunicaciones que sirve para el desarrollo de sistemas tutoriales inteligentes multimedia. De manera que, si el modelo se trasladara a otro dominio. Por ejemplo, en el dominio financiero, no tendría ninguna posibilidad de ser utilizado. Por otro lado, Guy et al., presentan un modelo semántico para la integración de datos aplicado al contenido multimedia en sistemas con protocolo de Internet por televisión de alto nivel [Guy et al., 2009]. Adicionalmente Fernández-Tobías et al., presentan un modelo evaluado en un el marco computación del trabajo que descubre los recorridos somáticamente entre los lugares de interés y la música de los artistas en un solo dominio [Fernández-Tobías et al., 2011]. Belk presenta un modelo multi-dominio para la personalización Web, basado en ontologías [Belk, Germanakos, & Tsianos, 2010]. Asimismo, [Ghani & Fano, 2002], presentan un modelo basado en el conocimiento para sitios de compras en línea en la Web. Hofmann describe un modelo basado en algoritmos para descubrir comunidades de usuarios y perfiles de interés [Hofmann, 2004].

Por otra parte, Mobasher presenta un algoritmo para crear una medida de similitud combinado y, poder así generar predicciones en la Web [Mobasher, Jin & Zhou, 2004]. Además, [Bedi, Kaur & Marwaha, 2007] presentan un modelo de trabajo a la perfección a través de diferentes versiones de ontologías en un domino turístico. Con el desarrollo de diversos modelos formales semánticos para los SR surgen nuevas formas para interpretar la información y un contenido que, a su vez, permite interactuar con los usuarios a través de los medios sociales en la Web.

2.3. Sistemas de Recomendación en los medios sociales

2.3.1. Introducción

Muchas veces, en nuestra vida cotidiana, se nos presentan situaciones en las que tenemos que tomar decisiones que a simple vista podemos considerar sencillas. La amplia información sobre los distintos productos, servicios y contenidos pueden hacer que un usuario se sienta desbordado al intentar discriminar sobre que producto, que servicio o que contenido cubre realmente sus necesidades. Ejemplos claros como elegir una película para ver [e.g., Miller, Albert, & Lam, 2003], un restaurante para cenar [e.g., McCarthy, 2002], un libro para leer [e.g., Huang et al., 2002], una canción para escuchar [e.g., Bu et al., 2010] o planear unas vacaciones [e.g., Mair, 2004]. Por lo que dichas decisiones pueden ser tareas bastante complicadas por una razón principal, la gran cantidad de información accesible sobre la música, libros, restaurantes, películas y la extensa información sobre destinos vacacionales que existen.

Por lo tanto, los distintos Sistemas de Recomendación han surgido para intentar disminuir las pequeñas o grandes dificultades sobre la toma de decisiones en nuestra vida cotidiana con múltiples opciones [Lü et al., 2012].

Actualmente, no podemos pasar por desapercibidos aquellos sitios de recomendación, de los cuales los medios sociales hacen uso para ofrecer secciones de tipo; “lo más leído”, “lo más votado” o “lo más enviado”, gustos y preferencias que nos ayudan a generar decisiones que finalmente llegan a ser parte natural de nuestro entorno social en el mundo y, que además, pueden hacer que un usuario se sienta desbordado al intentar discriminar sobre qué tipo de decisión debe realizar para cubrir sus necesidades.

En términos generales, los SR producen sugerencias (recomendaciones) sobre diferentes temas o productos (o acciones) dentro de un dominio en el cual está interesado

el usuario [Burke, Felfernig & Göker, 2011]. Por otro lado, los medios sociales permiten generar un contenido, que a su vez, es utilizado por los SR para proporcionar de distintas formas un tipo específico de información mediante técnicas de filtrado que intenta sugerir una información específica, tal como (blogs, noticias, música, planes de viaje, páginas web, imágenes, etiquetas, las comunidades de contenido, proyectos colaborativos, sitios de redes sociales, mundos virtuales de juegos, mundos virtuales sociales), dicha información que es susceptible de interesar a los usuarios [Ricci, Rokach, & Shapira, 2011].

Las modernas tecnologías de los medios sociales y el contenido de datos abundantes en la Web, hacen que compartir, colaborar con los demás y conectar entre sí, consienten una forma para crear una sociedad que cruza información de una manera más rápida, más fácil y más accesible a una comunidad exigente que crece desorbitadamente [Montaner, 2003].

Actualmente en la Web existen grandes colecciones de datos de medios heterogéneos desde distintos sitios sociales, bibliotecas digitales y aplicaciones de Internet [Shani & Gunawardana, 2011]. Además, muchos de los comportamientos sociales pueden reflejarse mediante la integración de estos datos en los medios heterogéneos, que a su vez, pueden ser usados para proporcionar recomendaciones en los medios de comunicación social [Chen, Zeng & Yuan, 2013].

En la Figura 2:8, se muestra las posibles relaciones que se puede utilizar para enriquecer los modelos existentes con los actuales datos que complementan diferentes dominios en los medios sociales.



Figura 2:8 Relaciones que permiten enriquecen las recomendaciones en los medios sociales.

Los Sistemas de Recomendación permiten alcanzar diferentes escenarios y contextos para lograr el acceso a la información [Jannach et al., 2010; R Burke, 2007]. Además, [Masters et al., 2008], comentan que los SR son herramientas para interactuar con grandes y complejos espacios de información.

Por otra parte, estos sistemas proporcionan una vista personalizada de distintas áreas de aplicación, dando prioridad a las cuestiones que puedan ser o no ser, de interés para el usuario [Konstan, 2004]. Por lo tanto, los SR son una ventaja al ser utilizados en los medios sociales, ya que mediante los SR es posible conocer los intereses de la información y contenido que, al mismo tiempo los usuarios valoran y opinan sobre ellos [Herlocker & Konstan, 2004]. Con los años y con el avance de las nuevas tecnologías que emergen continuamente, la definición de los SR ha ido evolucionado.

2.3.2. Evolución histórica

La sociedad se encuentra experimentando una rápida transformación en distintos aspectos gracias a las computadoras, el Internet y las redes sociales, de manera que la gente vive una vida social en Internet. Asimismo, el crecimiento de artículos disponibles en la Web convierte la información en grandes posibilidades para elegir un producto o servicio. Por lo tanto, es importante evaluar las distintas alternativas para discernir entre ellas a través de las nuevas tecnologías y aplicaciones ofrecidas en la Web. Los SR se consideran como una disciplina independiente desde mediados de los 90s [Foltz & Dumais, 1992; Stodolsky, 1995], cuando los investigadores comenzaron a utilizar las valoraciones de los usuarios sobre un contenido para ofrecer resultados a priori desconocidos para otros usuarios [Lü et al., 2012]. Es decir, que a todos los usuarios se les recomendaba el mismo conjunto de productos.

La descripción de la evolución de la IA y los SR aparecen en la década de los 90s, pero la IA se remonta hasta los años 40s, aunque no sería hasta 1950 cuando realmente consigue una verdadera repercusión en los sistemas [Turing, 1950]. Durante esa época Alan T., uno de los padres de la IA propone un test orientado a demostrar si una determinada máquina es inteligente o no [Turing, 2009]. En la década de los ochenta la IA vuelve otra vez a ponerse de moda, en esta ocasión de una forma más madura y precavida con la aparición de los Sistemas Expertos, más orientados a la aplicaciones prácticas que a la persecución abstracta de la inteligencia [López-Cuadrado, 2012]. Actualmente la IA y los SR en los medios sociales actúan en pleno desarrollo para generar y compartir información, así lograr sustituir las necesidades de los usuarios y empresas que demandan servicios [Zeng et al., 2010]. Algunos artículos relacionados se citan a lo largo del texto.

Los Sistemas de Recomendación Demográficos fueron los primeros SR que se comenzaron a utilizar. En este tipo de sistemas, el usuario es clasificado en un perfil según sus características demográficas y las recomendaciones se calculaban basándose en la información de necesidades, preferencias o gustos de dicho perfil demográfico. Grundy en su obra presenta el primer sistema con estas características [E Rich, 1979]. No obstante, en la actualidad continúa su éxito. Por ejemplo Beel y Langer presentan un trabajo sobre

la importancia de considerar la demografía y otras características para recomendaciones entre los usuarios [Beel & Langer, 2013].

Más adelante se inicia con la utilización de la recuperación de la información personalizada, dicha información intenta refinar la búsqueda de documentos incorporando información sobre las necesidades específicas del usuario. De esta manera, una consulta puede devolver una recomendación distinta, dependiendo del usuario que la formule, ya que filtra la información de forma diferente según el perfil de un usuario [Sheth, 1994]. Un tipo de sistemas que también se utilizaba en los inicios de los SR son los Sistemas Basados en Contenido, lo cuales contienen información detallada sobre el contenido de los productos, como sus atributos o descripciones de los mismos. Además, el perfil de usuario se calcula utilizando la información de contenido de los productos que son de su preferencia. Uno de los primeros SRBC es el sistema propuesto por Lang sobre las preferencias de un usuario en un sistema de filtrado de noticias [Lang, 1995]. Posteriormente surgieron los Sistemas de Recomendación Colaborativos [Goldberg et al., 1992]. Finalmente, los Sistemas de Recomendación Basado en Conocimiento [Burke, Hammond & Yound, 1997], dichos sistemas calculan las recomendaciones intentando hacer inferencias sobre las necesidades y preferencias de los usuarios en el momento de la recomendación. Burke presenta uno de los primeros sistemas que ayuda a los usuarios a encontrar adecuados apartamentos de alquiler en Chicago [Burke et al., 1996].

En la actualidad los SR han evolucionado y es posible encontrarlos en diversos ámbitos de aplicación, ya que permiten la consecución de diferentes escenarios y ambientes para lograr el acceso a los contenidos, además, son herramientas para interactuar con los grandes y complejos espacios de información [Rosaci & Sarné, 2010 ; Hernández del Olmo & Gaudioso, 2008]. Los SR se utilizan para recomendar elementos potencialmente interesantes para los usuarios en diferentes dominios. Por lo tanto, existen para satisfacer la necesidad de prestar asistencia potencialmente interesante para los usuarios en múltiples dominios en los medios sociales [Stavrianou & Brun, 2013]. Por ejemplo, en el comercio electrónico, donde se han convertido en una herramienta fundamental para los proveedores en línea [e.g., Schafer, Konstan, & Riedl, 2001], el turismo [e.g. Garcia, Sebastia, & Onaindia, 2011; Borràs, Moreno, & Valls, 2014], películas [e.g., Golbeck & Hendler, 2006], música [e.g., Hyung, Lee, & Lee, 2013], restaurantes [e.g.,

Hung-Wen Tung & Von-Wun Soo, 2004], programas de televisión [e.g., Martinez et al., 2009], juegos sociales [e.g., Dugan et al., 2007], mundos virtuales [e.g., J. F. McCarthy, 2001], en las finanzas [e.g., Alejandro Rodríguez et al., 2012], la salud [e.g., Casado-Lumbreras, 2012], por mencionar solo algunos de ellos.

No obstante, en cada dominio se presentan diferentes problemas a los que hay que facilitar diferentes soluciones. Por lo tanto, la necesaria capacidad de evolución sobre los distintos productos y la información de los usuarios han implicado que los Sistemas de Recomendación se hayan diversificado en gran manera. De manera que, los SR sirven para generar recomendaciones, usan las entradas del usuario activo, pero también información sobre los ítems o información del resto de usuarios del sistema, que actúan como colaboradores. En este sentido, la realimentación por parte de los usuarios es muy importante de cara a albergar una información más completa ante futuros procesos de generación de recomendaciones. La Figura 2:9, refleja el proceso genérico para el desarrollo de las recomendaciones en un sistema.

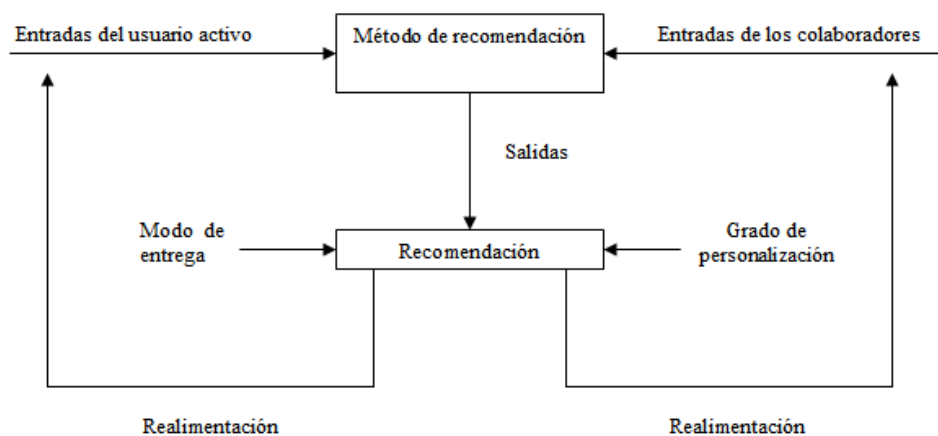


Figura 2:9 Proceso genérico para la generación de recomendaciones en un Sistema de Recomendación.

Para realizar una recomendación a un usuario, es necesario conocer algún tipo de información sobre sus preferencias o gustos. Además, dependiendo del tipo de sistema

también es necesaria la información sobre los temas a recomendar o la información reunida sobre el resto de usuarios del sistema (i.e., comunidad de usuarios o colaboradores). Dicha información es suficiente para realizar las recomendaciones pertinentes que además constituyen la entrada o entradas de un SR.

Existen distintas formas para recomendar, a continuación se dan algunos ejemplos que han servido de base para su estudio. Terveen et al., presentan “PHOAKS”, un sistema experimental para solucionar el problema de encontrar información relevante y de alta calidad en la Web, usando el enfoque colaborativo en el que los usuarios recomiendan distintos temas a otros usuarios. El sistema trabaja reconociendo, concordando y redistribuyendo automáticamente recomendaciones de recursos Web extraídos de mensajes de noticias [Terveen, Hill, & Amento, 1997].

Por otro lado, Balabanović & Shoham presentan un sistema llamado “FAB” orientado a la recomendación de “URL” (*Uniform Resource Locator*) que combina el uso de información por extensión con el enfoque colaborativo [Balabanović & Shoham, 1997]. Resnick & Varian presentan un trabajo sobre un sistema llamado “SITESEER”, dicho sistema recomienda páginas Web relevantes y usa las listas de favoritos y la organización de registros como una declaración implícita de intereses respecto al contenido subyacente, el sistema mide el grado de solapamiento con las de otros usuarios [Resnick & Varian, 1997].

En la actualidad los SR que se desarrollan varían desde la manera en la que las sugerencias son entregadas a los usuarios, la experiencia con la que se construyen y la manera en la que se presentan [Hsieh, 2011]. La diferencia se encuentra en la manera de como interaccionan con las características de los usuarios, productos o servicios que son de interés para los usuarios. Como resultado, los SR son una herramienta de gran utilidad para ayudar a los usuarios en sus procesos de búsqueda y recuperación de la información almacenada en la Web desde diferentes paradigmas.

2.3.3. Características de modelado de perfiles de usuarios

Los SR necesita tener información sobre el usuario al que se pretende recomendar. Para ello, es necesario calcular un perfil de usuario a partir de la información que se conoce sobre él. Para esta tarea se han aplicado distintos enfoques, como el modelado del comportamiento [Eckert, Levin, & Pieraccini, 1997], modelado de intereses o de intenciones [W. Wang et al., 2013]. Con esta información se pueden realizar clasificaciones de los usuarios, agruparlos (en inglés, clustering), extraer patrones de comportamiento y realizar predicciones sobre los mismos, acordado por [Canberk, Akyildiz, & Oktug, 2011]. Por lo tanto, un perfil es una colección de propiedades de un objeto. Cuando se hace un perfil de un usuario en línea, estas propiedades incluyen, por ejemplo, el método de conexión utilizado, el terminal utilizado, patrones de comportamiento e intereses del usuario.

Cuando se hace el perfil de un cliente, las propiedades normalmente preferidas incluyen la edad, la residencia habitual, el nivel del salario, así como las distintas formas de contacto (i.e., email, teléfonos - fijo o móvil, etc.), hábitos de compra, etc. También se hacen perfiles de documentos. En este caso las propiedades más importantes suelen ser el nombre del documento, el autor, la fecha de creación y las palabras. Por otro lado, si se habla del perfil de un producto, las características incluyen, por ejemplo, el productor, precio e información técnica del mismo.

Además, los modelos de perfiles de usuarios permiten realizar una personalización para un gran número de usuarios. En consecuencia de dividir y seleccionar la información para cada usuario utilizando para ello tecnologías de IA, como son: redes neuronales, redes bayesianas, lógica borrosa y razonamiento basado en casos. También se suelen emplear heurísticas basadas en el dominio de aplicación para mejorar el desempeño de los sistemas. Por otro lado, Cruz, Peñalvo y Romero comentan que la personalización de contenidos es cada vez más utilizada como medio de agilizar la entrega de una información hacia el usuario, convirtiendo al sistema en más útil y atractivo, de forma que el usuario se sienta estimulado al utilizarlo [J. Liu, Dolan, & Pedersen, 2010 ; Cruz, Peñalvo, and Romero

2003]. De tal forma que la personalización se viene convirtiendo en un requisito esencial en los Sistemas de Recomendación y en el modelado de perfiles de usuarios de hoy en día. Por ejemplo Bouno et al., presentan técnicas para analizar los datos de perfiles en la Web y generar recomendaciones en el dominio del comercio electrónico [Buono et al., 2002].

Los modelos de perfiles de usuarios son usados para representar cosas del mundo real, por lo tanto, necesitan saber qué características son necesarias para modelar el contenido. En ocasiones, los modelos hacen uso de tecnologías para la representación del contenido, estos son dados a través de las ontologías, tal y como se describen en la Sección 2.1.5. Por lo tanto, la especificación formal explícita de cómo representar objetos, conceptos, relaciones y otras entidades que se asume que existen en una determinada área o dominio de interés, los modelos de perfiles, en consecuencia proyecta características del mundo real, hacia al mundo definido a través de las ontologías. Por ejemplo, Mobasher presenta dos métodos para la construcción de perfiles, con base en datos históricos de los usuarios anónimos, PACT (*“Profile Aggregations based on Clustering Transactions”*) y ARHP (*“Association Rule Hypergraph Partitioning”*), [Mobasher et al., 2002].

2.3.4. Características de modelado de productos

Cuando las personas debemos realizar una elección sobre un producto o servicio del que no tenemos conocimiento previo de su calidad o prestaciones, solemos recurrir a la experiencia o consejos de otros. En consecuencia, buscamos a personas conocidas familiarizadas con los temas que perseguimos para que nos recomienden uno de entre las diferentes posibilidades.

Por lo tanto, es necesario conocer las características de un producto o servicio que permita a los usuarios llevar a cabo acciones, aproximaciones sobre sus preferencias y gustos hacia aquellos dominios de su interés. No obstante, Ardissono et al., comentan

la importancia de la generación de descripciones sobre los distintos productos almacenados en la Web [Ardissono et al., 1999; Brusilovsky, 2001].

Los sistemas de información necesitan conocer los intereses de los usuarios para llegar a realizar sus funciones de forma efectiva [Crabtree & Soltysiak, 1998]. Pero, como podemos describir un producto utilizando su estructura, esto es, la descripción del producto que se expresa como un conjunto de categorías agrupadas por medio de una jerarquía [Kwak et al., 2010]. Estas tareas se realizan de formas distintas según el tipo de SR que se esté desarrollando, incluso algunos sistemas no realizan modelado de productos o de usuarios, ver la información de la Sección 2.2.1. Para conocer los productos que existen en el sistema se calcula también un perfil. Estos perfiles de productos se pueden utilizar para analizar los productos, realizar una extracción de variables semánticas que permiten analizar la diversidad de los mismos.

2.3.5. Características de filtrado de información

Con el volumen de información que existe en la Web semántica y, además crece a un ritmo exponencial, las necesidades de los usuarios para la búsqueda de información, resultan generar filtrados con eficacia [Boardman & Lu, 2007], accesibles y fáciles de manejar [Abel, Hauff, & Houben, 2012]. Por lo tanto, es necesario resolver los desafíos para la búsqueda y análisis de la información sobre los distintos productos y servicios que permitan satisfacer las necesidades de los usuarios en un mundo real.

Existen herramientas y tecnologías que permiten analizar el filtrado de información, tales como, agentes inteligentes de información en la Web [e.g., Bradley, 1999], motores de búsqueda [e.g., Chu & Rosenthal, 1996], procesamiento de lenguaje natural [e.g., Vázquez et al. 2009], ontologías [e.g., Fensel & Harmelen, 2001], redes neuronales [e.g., Rodriguez-Gonzalez et al., 2010], los sistemas expertos [e.g., Duan, Edwards, & Xu, 2005]. Además, estudios recientes muestran que dichas tecnologías y herramientas hacen uso de los datos almacenados en la Web semántica para ayudar a

detectar incidentes sobre temas específicos [e.g., Weng & Lee, 2011; e.g., Clark & Roos, 2014; e.g., Saha Roy et al., 2014; Vázquez-Ramírez, 2014].

La capacidad de poner el conocimiento en la Web, que permite compartirlo y reutilizarlo a través de mecanismos estándares, proporciona nuevos retos interesantes en el área de la IA [J. A. Hendler, 2005]. Por lo tanto, una vez conocidos los perfiles de usuarios y la búsqueda de los productos sobre un dominio, es importante conocer las características al realizar un filtrado de información basándose en cómo es de adecuado un producto hacia un usuario en concreto y, que a su vez, permita predecir la valoración que un usuario daría a un producto. Para ello existen distintos enfoques, como el filtrado basado en las características de los productos, basado en reglas o en filtrado colaborativo.

Existen distintos sistemas para el filtrado de información dependiendo de los intereses de los usuarios [Rodríguez-González et al., 2012]. Por un lado, tenemos los filtros basados en el contenido mediante los cuales, el sistema de forma automática trata de recomendar productos que el usuario ya había adquirido o consumido en el pasado comparándolo con su perfil [Golbeck & Mannes, 2006]. Mientras que el filtrado colaborativo mostraría al usuario productos que han gustado a usuarios con perfiles similares al suyo y le sugeriría productos que otros ya han adquirido [Massa & Avesani, 2004]. Finalmente, tendríamos sistemas de filtrado de la información social, que se basaría en la recomendación de informaciones y productos mediante las valoraciones voluntarias y colaborativa por parte de los usuarios que están dados de alta en dichos sistemas [Lerman, 2006]. Para una mejor comprensión de los conceptos y el tema se sugiere ver la Sección 2.2.

Por otro lado, Agichtein et al., comentan una forma más concreta utilizada dentro del filtrado basado en el contenido, cuando un usuario encuentra interesante un documento, un servicio o un producto, el sistema puede facilitarle contenidos similares dependiendo de los pesos que se asignan a las palabras que mejor describen el contenido del texto o las características de un producto [Agichtein et al., 2008]. Asimismo, el contenido se compararía con los pesos que se guardan dentro del perfil del usuario y se enfrentaría a otros temas. A este proceso se le conoce como “feedback por relevancia o retroalimentación por relevancia” [Kosak & Lang, 1999]. Cabe mencionar que estos

sistemas son eficaces cuando las recomendaciones se realizan basándose tan sólo en los pesos que se almacenan dentro del perfil del usuario y basándose en las puntuaciones que el usuario ha dado a los distintos contenidos que ha consultado en el pasado.

Aproximaciones respecto al filtrado de información han sido desarrolladas por múltiples servicios cuyo objetivo es ofrecer a sus clientes productos eficientes a sus necesidades que además se encuentren relacionados a sus compras anteriores o a los ítems que habían sido consultados. Asimismo poner en contacto a clientes con gustos similares. Sin embargo, la Web semántica ofrece una vasta información para realizar filtrados a través de modelos, técnicas y aplicaciones que son usados por los SR que basan el contenido en las sugerencias que hace un usuario a sus amistades sobre productos o servicios que encuentra interesantes.

En este sentido, la presente investigación se desarrolla con herramientas y tecnologías semánticas que nacen a partir de un enfoque analizado, ya que dicho enfoque permite analizar el filtrado de información basado en el contenido y el conocimiento. Por lo tanto, de manera siguiente se incluye una sección del filtrado basado en el contenido que permite ayudar a entender los conceptos relacionados con la técnica.

2.3.6. Recomendación basada en contenido

La abundancia de información en la Web ha establecido un rápido crecimiento a la necesidad de encontrar lo que queremos cuando lo necesitamos y de la manera que mejor se adapte a nuestras necesidades. Como consecuencia, el usuario necesita el apoyo para obtener una información personalizada de acuerdo a sus gustos y preferencias en una amplia información que se almacena en la Web.

Los medios de comunicación social han supuesto una revolución en Internet, modificando la forma en que los usuarios gestionan y acceden a la información de manera casi inmediata y prácticamente de forma universal. Por lo tanto, muchas fuentes de información incorporan SR como una forma de personalizar sus contenidos para diferentes dominios. Dichos sistemas tienen el efecto de guiar a los usuarios en forma personalizada

a los objetos interesantes en un gran espacio de posibles opciones [Monfil-Contreras et al., 2012]. Además, los Sistemas de Recomendación basados en contenido tratan de recomendar artículos similares a los que un usuario le ha gustado en el pasado [Lops, Gemmis, & Semeraro, 2009]. De hecho, el proceso básico consiste en hacer coincidir los atributos de un perfil de usuario en el que las preferencias y los intereses se almacenan, con los atributos de un establecido objeto de contenido (i.e., tema, artículos, productos, características), esto es con el fin de recomendar a los usuarios nuevos productos interesantes [Lops Pasquale, 2011]. Además, implementan un enfoque para analizar los documentos y/o descripciones de elementos previamente valorados por un usuario. Por lo tanto, construyen un perfil de intereses para el usuario sobre la base de las características de los objetos valorados por ese mismo usuario [Mladenec, 1999]. Por ejemplo, Leis et al., presentan un estudio para conocer las características de contenido sobre grupos de Facebook relacionados con la alimentación saludable [Leis et al., 2013]. En la Figura 2:10 se muestra una arquitectura de alto nivel para el proceso básico de recomendación basada en contenido.

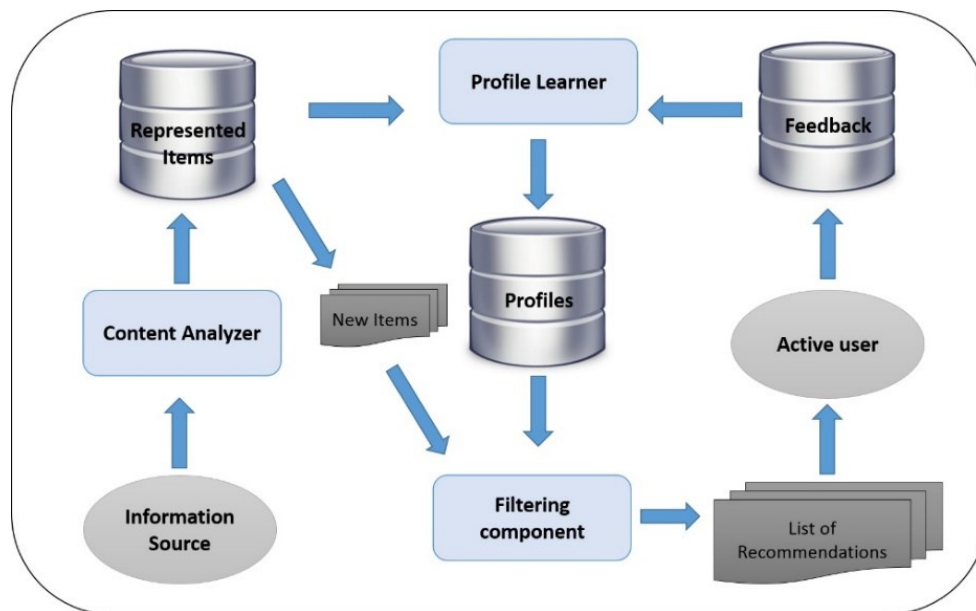


Figura 2:10 Arquitectura de alto nivel para un Sistema de Recomendación basado en contenido según [Valdéz 2012].

El proceso de recomendación se realiza en tres pasos. Cada paso es manejado por un componente separado. En primer lugar, el analizador de contenido representa el contenido de los productos (i.e., documentos, páginas web, noticias, descripciones de productos, sus características, etc.), dicho contenidos son procedentes de fuentes de información en una forma adecuada para los siguientes pasos de procesamiento. En segundo lugar, los elementos de datos son analizados por técnicas de extracción de características con el fin de desplazar el elemento de la representación de espacio de la información original a la del destino. El módulo de "perfil de aprendizaje" recoge datos representativos de las preferencias de los usuarios o retroalimentación (Feedback) y trata de generalizar estos datos, con el fin de construir el perfil de usuario. El componente de filtrado explota el perfil de usuario para sugerir temas relevantes, haciendo coincidir la representación del perfil frente a la de los elementos para ser recomendados. Así, el contenido de las descripciones del artículo procedentes de las fuentes de información es procesado por el "analizador de contenido". No obstante, los procesos de recomendación basados en contenido manifiestan ventajas y desventajas. A continuación se mencionan algunas de las ventajas más destacadas para el proceso:

- **Independencia de usuario** – Los Sistemas de Recomendación basados en contenido explotan únicamente calificaciones proporcionadas por el usuario activo para construir su propio perfil.
- **Transparencia** – Los sistemas basados en contenido proporcionan explícitamente funciones de contenido o descripciones de un elemento que son causas para una recomendación. Las características son indicadores para consultar la recomendación.
- **Nuevo artículo** – Los sistemas basados en el contenido son capaces de recomendar artículos que todavía no han sido calificados por otros usuarios.

Para las desventajas en los procesos de recomendación basada en contenido podemos destacar las siguientes:

- **Análisis de contenido limitado** – El proceso de recomendación presenta técnicas que tienen un límite natural en el número y tipo de características que están asociadas, ya sea de forma automática o manualmente, con los objetos que ellos recomiendan. Con frecuencia se necesita el conocimiento de un dominio, por ejemplo, para obtener recomendaciones de películas el sistema necesita conocer los actores y directores, y, a veces, también se necesita hacer uso de otros recursos como las ontologías de dominio.
- **Especialización** – Los SR basados en contenido no tienen ningún método inherente para encontrar algo inesperado.
- **Nuevo usuario** – Los SR basados en contenido tienen que recoger calificaciones suficientes antes de entender las preferencias del usuario y proporcionar una información precisa para las recomendaciones.

Por otra parte, las técnicas basadas en el contenido en los SR tienen una descripción de cada elemento disponible que se encuentra almacenado en la base de conocimiento. De manera que el sistema compara las descripciones de los usuarios frente a las descripciones del elemento almacenado y recomienda los elementos que mejor se adapten a las necesidades. Además, esta técnica consiste en usar las características de los productos adquiridos y valorados por el usuario para encontrar otros similares que también le puedan gustar. Estos sistemas necesitan técnicas apropiadas para la representación de los elementos y la producción del perfil de usuario, y algunas estrategias para comparar el perfil de usuario con la representación de un artículo [Koren, Bell, & Volinsky, 2009]. Por lo tanto, la técnica de filtrado basado en contenido en los Sistemas de Recomendación explota el contenido de elementos de datos para predecir su relevancia en función del perfil del usuario [Aurnhammer, Hanappe, & Steels, 2006].

Actualmente la recuperación de la información se está aplicando en una variedad de dominios de aplicación en la Web. Además, es una disciplina que involucra la localización de una información dentro de un almacén de información o base de datos [Varelas & Voutsakis, 2005]. Por otro lado, los motores de búsqueda se basan en palabras claves para proporcionar una respuesta a las necesidades de los usuarios en los significados de un contenido, la búsqueda semántica de contenidos es de gran importancia para los investigadores en el campo de la recuperación de la información [Fernández et al., 2011]. Por lo tanto, en esta investigación se hace uso de un motor de búsqueda para la extracción del contenido que permita encontrar respuestas a las necesidades de un usuario. En la siguiente sección se describen las técnicas y aplicaciones de los modelos semánticos.

2.4. Técnicas y aplicaciones para modelos semánticos en medios sociales

Los SR son un campo amplio en la IA cuyos datos se adecuan a la aplicación de distintas técnicas de este tipo, ya que existen diferentes modelos semánticos aplicados al dominio de los medios sociales con gran cantidad de productos y servicios, como puede ser la recomendación de libros, películas, canciones, servicios de comida, viajes, etc. El análisis de los medios sociales en la utilización de modelos para los SR es amplia, por lo que en la presente tesis doctoral, se ha limitado a presentar modelos basados en ontologías y modelos basados en Sistemas Expertos, ya que dichos modelos contienen componentes principales en valores singulares que han sido aplicados a los SR con éxito, por lo que serán descritos en las siguientes secciones.

2.4.1. Modelos para la representación basados en ontologías

Los Sistemas de Recomendación hacen uso esencialmente de modelos en conocimiento que capturan información y representan dicho conocimiento sobre ciertos dominios [Viinikkala, 2003]. Por lo tanto, las ontologías pueden proveer los mecanismos para organizar y almacenar los componentes genéricos de los SR que incluyen esquemas de las bases de datos, objetos de interfaz de usuario y programas de aplicación. Es decir, las ontologías constituyen un nuevo enfoque en la investigación y desarrollo de la disciplina de los SR, acordado por [Pisanelli, Gangemi, & Steve, 2002; Guarino, 1998].

Las ontologías están desarrollándose y aplicándose a una extensa variedad de áreas de aplicación emergentes, tales como modelado de empresas, diagnósticos, toma de decisión, planeamiento y adaptación, modelado de procesos y sistemas [Obitko, 2001]. Además, son una de las formas más útiles para representar las variantes terminológicas con relación a un modelo conceptual [Cea & Montiel-Ponsoda, 2012]. Los modelos para la representación del conocimiento basados en ontologías son usados en distintos dominios. Por ejemplo, en el apoyo a la toma de decisiones mediante un modelo ontológico presentado por [Borges, Corniel, & Gil, 2009], dicho modelo utiliza la información de la Web para ofrecer recomendaciones personalizadas.

Las ontologías tiene como objetivo proporcionar un lenguaje que puede ser utilizado para describir las clases y las relaciones entre ellos que son inherentes a documentos y aplicaciones en la Web [Á García-Crespo & Rodríguez, 2010]. Además, las ontologías proporcionan potencialmente un mecanismo bien fundamentado para la representación y el intercambio de la información estructurada [Ye & Coyle, 2007]. Por ejemplo, Dey et al., presentan un modelo basado en ontologías para sistemas de contexto informal [Dey, Abowd, & Salber, 2001], otro trabajo presentado por [Kindberg, Barton, & Morgan 2002], proponen un modelo en el que cada objeto tiene una descripción en la Web.

Por otro lado, existen tres tipos de ontologías, a los que hay que sumar las ontologías creadas para una tarea específica, como por ejemplo, el diagnóstico de una enfermedad [Steve et al., 1998].

- Ontologías de un dominio; que representan el conocimiento especializado de un determinado campo, como la medicina.
- Ontologías genéricas; en las que se representan conceptos generales.
- Ontologías representacionales o meta-ontologías; en las que se conceptualizan los formalismos de representación del conocimiento.

Asimismo, las ontologías tienen una amplia diversidad de aplicación en distintos dominios, tales como, la biomedicina [Bates & Cohen, 2001]. Por ejemplo, Smith et al., presentan ontologías medicas a partir de historias clínicas electrónicas [e.g., Smith et al., 2005]. en el dominio financiero [e.g., Zhang, Zhang & Ong, 2000], negocios inteligentes [e.g., Vela & Declerck, 2009], entretenimiento [e.g., Carbonaro & Ferrini, 2007], de servicios electrónicos [e.g., Chang, Dillon, & Hussain, 2007], las redes sociales [e.g., Jung & Euzenat, 2007], por citar algunos de ellos.

A principios de la década de los 90's, se creó un conjunto de lenguajes para la implementación de ontologías basadas en IA. Básicamente, el paradigma para la representación del conocimiento [MacGregor & Bates, 1987]. No obstante, actualmente existen leguajes desarrollados para describir las ontologías, pero no fue hasta el año 2004 que la W3C estandarizo y formalizo un Lenguaje de Ontologías Web (OWL), el lenguaje permite definir clases, propiedades, individuos y hacer valer las propiedades de los conceptos que permiten generar una semántica formal, ver la Sección 2.1.5. No obstante, a continuación se menciona una breve descripción acerca de los lenguajes más relevantes usados para el desarrollo de modelos basados en ontologías de la Web semántica.

El Formato KIF (*Knowledge Interchange Format*) por sus siglas en inglés, es un lenguaje basado en la lógica de primer orden y creado como un formato de intercambio para los diferentes sistemas de representación del conocimiento [Genesereth & Fikes, 1992; Ginsberg & Geddis, 1991]. Ontolingua es un sistema para la descripción de ontologías en un formato canónico que le permite traducirlas fácilmente en una variedad

de sistemas para la representación del conocimiento y razonamiento [Gruber, 1992]. Por otra parte, [MacGregor & Bates 1987], presentan LOOM, un lenguaje que ofrece una clasificación automática de conceptos y está basado en lógicas descriptivas y normas de producción para la representación del conocimiento. El lenguaje OCML (*Operational Conceptual Modelling Language*) sirve de apoyo para el modelado a nivel de conocimiento. En la práctica, este rol implica que OCML se centra en la lógica en lugar de las primitivas del nivel de aplicación [Motta, 1998]. Balaban comenta sobre F-Logic (*Frame Logic*) un marco de referencia base para los lenguajes de descripción del conocimiento y, es utilizado para proporcionar una relación completa de este tipo de lenguajes sin perder la semántica directa y su carácter descriptivo [Balaban 1995].

Por lo tanto, los distintos lenguajes permiten el desarrollo de modelos ontológicos que pueden ser usados por SR aplicados a los medios sociales en la Web. En consecuencia, es necesario continuar con el desarrollo de modelos conceptuales y la mejora de los ya existentes, que permitan dar solución a las distintas necesidades de los usuarios y que además, sirvan en el modelado de contenido en los medios sociales.

2.4.2. Modelos para la representación en Sistemas Basados en Conocimiento

Los Sistemas Basados en Conocimiento son aquellos que utilizan el conocimiento de un experto en un dominio para representar un conocimiento, también conocidos como sistemas expertos [Zha, Du, & Qiu, 2001], representan un paso delante de los sistemas de información convencionales al pretender representar funciones cognitivas del ser humano como el aprendizaje y el razonamiento. Sin embargo, el termino sistema experto se aplica frecuentemente a cualquier sistema que utiliza tecnología de sistemas expertos (i.e., lenguajes específicos, programas o incluso hardware diseñado para ayudar al desarrollo y ejecución de sistemas expertos), [J Giarratano & Riley, 1998]. El objetivo de estos sistemas es ayudar a la creación, transferencia y aplicación del conocimiento en diversos dominios [Alavi and Leidner 2001], que además soportan actividades de toma decisiones [G. Klein & Orasanu, 1993]. Por lo tanto, la información que ofrecen dichos sistemas es

útil de identificar, de resolver problemas y tomar decisiones. Además, aportan importantes ventajas competitivas para un conocimiento que es necesario gestionarlo adecuadamente.

Generalmente los sistemas expertos basados en conocimiento se diseñan de forma diferente a los sistemas convencionales, debido a que, en general, los sistemas expertos tratan de solucionar problemas que no tienen una solución algorítmica [Chan, 2005].

En los últimos años, las redes sociales en la Web se han convertido en un área de negocios, de comunicación y entretenimiento junto a la evolución de las nuevas tecnologías, de tal forma que los SR y los sistemas expertos, aportando modelos para la representación del conocimiento cada vez más efectivos. Por ejemplo, en problemas industriales. Hunter et al., presentan un modelo de representación para un proceso de inspección industrial, creando un sistema que permite automatizar la planificación y ejecución de actividades de inspección [e.g., Hunter et al., 2007]. Asimismo, en problemas comerciales [e.g., Ngai & Wat, 2005], de diagnósticos médicos [e.g., Rocha, & Oliveira, 1997], finanzas [e.g., Wagner, Otto, & Chung, 2002], vigilancia y control de procesos [e.g., Agogino, Srinivas, & Schneider, 1988], negocios inteligentes [e.g., Eom, 1996], diseño y fabricación [e.g., López-Cuadrado, 2012].

Los modelos para la representación del conocimiento se diseñan de tal forma que se diferencian de los modelos que se emplean en los sistemas convencionales, ya que los modelos basados en conocimiento, lo reitero, tratan de solucionar problemas que no tienen una solución algorítmica [J Giarratano & Riley, 1998]. Por otro lado, López Cuadrado presenta una arquitectura básica de un Sistema Basado en Conocimiento, en ella se puede observar el proceso genérico que realizan los sistemas expertos (ver Figura 2:11).

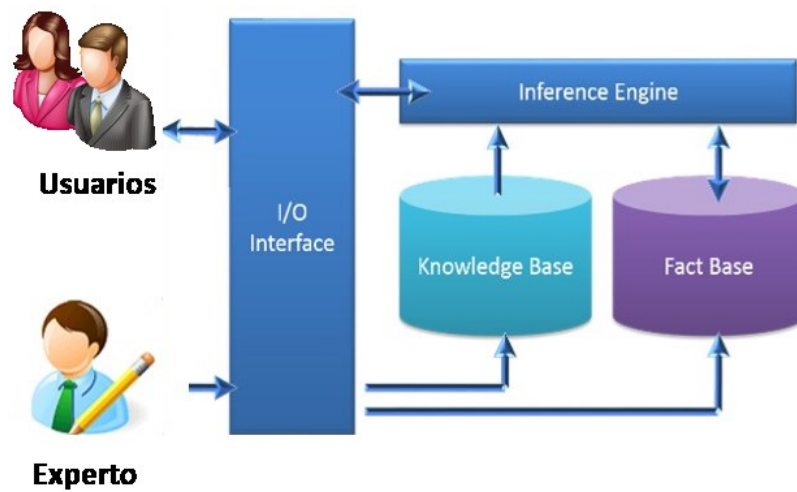


Figura 2:11 Arquitectura general de un sistema experto [e.g., Lopez-Cuadrado, 2012].

La arquitectura consta principalmente de una interfaz de entrada/salida, la interfaz permite que el usuario pueda enviar datos y recibir las respuestas del sistema y, además el experto pueda introducir sus conocimientos. Asimismo, consta de una base de hechos en la que se almacena la información que se recibe del entorno. Una base de conocimiento, en la que se encuentran las reglas que un experto en el dominio del problema utiliza en la toma de decisiones. Un motor de inferencia, para elegir la decisión a tomar partiendo de las reglas y los conocimientos adquiridos. Finalmente de un módulo de explicación de las decisiones que ha tomado.

Las técnicas para los modelos en la representación del conocimiento empleados con más frecuencia en el desarrollo de sistemas expertos son mencionando brevemente a continuación.

Lógica Borrosa: Una técnica que se basa en representar el conocimiento utilizando variables lingüísticas de vagas definiciones, similares a las que utiliza un experto en su trabajo diario, expresándolas en forma de conjuntos y funciones de pertenencia [Zadeh, 1983].

Objetos: son una colección de información que representa una entidad perteneciente al mundo real y una descripción de cómo esta información es manipulada [O’Keefe, Balci, & Smith, 1986].

Frames o marcos: representan el conocimiento relacionado con un problema determinado. Contienen un conjunto de pares atributo-valor que representan una situación, un objeto o un proceso del mundo real [Gonzalez & Dankel, 1993].

Lógica de predicados: se basa en proposiciones que expresan relaciones entre objetos así como cualidades y atributos de dichos objetos., éstas relaciones entre atributos se denominan predicados [Nosek & Roth, 1990].

Reglas: son una forma natural de expresar el conocimiento, tienen una estructura del tipo If <<condición>> then <<acción>>, donde el lado izquierdo de la regla se denomina antecedente o premisa, y el lado derecho se denomina consecuente o conclusión [Gonzalez & Dankel, 1993].

Redes semánticas: la técnica representa objetos, conceptos o situaciones en forma de nodos y relaciones entre ellos [JC Giarratano & Riley, 1989].

Redes Bayesianas: La técnica consisten en aprovechar las relaciones de dependencia (y por tanto también las de independencia) existentes entre las variables de un problema antes de especificar y calcular con los valores numéricos de las probabilidades entre un subconjunto de atributos [Díez et al., 1997].

Actualmente, distintos tipos de técnicas se desarrollan para el razonamiento en la creación de Sistemas Basados en Conocimiento, por ejemplo, las redes de neuronas y los sistemas basados en casos. No obstante, autores comentan que el razonamiento en dichas técnicas es un tipo específico de reglas.

2.4.3. Lógica borrosa

Lógica borrosa es una herramienta para modelar información desde la Web y poder inducir nuevo conocimiento [Stoilos et al., 2005]. Cuando tratamos de modelar conceptos del mundo real, muchas veces necesitamos simplificarlos para que su expresión en el lenguaje de la lógica sea sencilla. Estas simplificaciones provocan una pérdida de información que en algunos casos puede no ser asumible. Además, la lógica borrosa es una de las soluciones que se presentan para hacer que la lógica tenga grados de verdad mediante la incorporación de probabilidades. De manera que una sentencia deja de ser verdadera o falsa para ser verdadera en un cierto grado, es decir se cambian $\{0, 1\}$ por el intervalo $[0,1]$. El empleo de lógica difusa en la Web semántica puede hacer que la definición de las ontologías sea más precisa y más robusta. En este sentido también permite que los usuarios de la Web semántica puedan expresar conceptos relativos como el indicarles a los mismos el grado de certeza que tiene en los resultados [Stoilos, Stamou, & Pan, 2010].

Varias aplicaciones modernas hacen uso de las tecnologías semánticas con el fin de mejorar su desempeño y la gestión del conocimiento en la Web. Actualmente existen diversas propuestas para tratar la información y el conocimiento semántico con lógica borrosa [Straccia, 2006]. Por ejemplo, los algoritmos de razonamiento destinados a proporcionar inferencia en apoyo a la información vaga [Sanchez, 2006], las teorías de conjuntos difusos [J & Bo, 2008] a partir de descripciones lógicas basadas en lenguajes de representación del conocimiento [Baader, 2003] y, las ontologías difusas [Widyantoro & Yen, 2001], por mencionar solo algunas de ellas.

En este sentido, las ontologías difusas básicamente son una estructura que pueden ser definidas como un conjunto de conceptos, de relaciones difusas entre los conceptos, de una jerarquía de conceptos y taxonomías, además de relaciones no jerárquicas asociadas entre ellos [Stoilos et al., 2005]. Por ejemplo, Aquin, Lieber, & Napoli presentan un sistema para la creación de un portal semántico basado en ontologías y datos difusos [Aquin, Lieber, & Napoli, 2006]. Bobillo & Straccia, presentan un enfoque mediante la identificación de las diferencias sintácticas sobre el lenguaje de ontologías

difusas utilizando OWL [Bobillo & Straccia, 2011]. C.S. Lee, Jian & Huang, presentan una ontología difusa mediante conceptos nítidos para noticias de resumen en una página Web [C.S. Lee, Jian, & Huang, 2005]. Además, el uso de modelos difusos en lenguaje OWL en la representación del razonamiento para sistemas inteligentes asistido por ordenadores, proporcionan un apoyo eficaz a los usuarios [Ma et al., 2011].

Por lo tanto, la lógica borrosa y las ontologías tienen un papel importante, representan un medio para el intercambio de conocimientos entre usuarios y sistemas [Parry, 2006]. Además, son una herramienta importante para la recuperación de la información almacenada en la Web. En la presente tesis doctoral se hace uso de la lógica borrosa a partir de componentes y conjuntos difusos que proporcionan una representación natural para el conocimiento humano sobre un dominio, además de transformar el contenido a un mundo real para su fácil comprensión incluso para usuarios no expertos.

2.4.4. Transformación de componentes y conjuntos difusos

Hoy en día, los sistemas basados en reglas difusas se aplican con éxito en una amplia gama de problemas del mundo real para distintos dominios [Hadavandi, Shavandi, & Ghanbari, 2010] y, además en muchas aplicaciones del mundo real [Le, Nguyen, & Swiatek, 2010; Li & Chan, 2011; Qin et al., 2011]. Aunque en un sentido general un sistema puede ser definido matemáticamente, actualmente los ingenieros todavía prefieren optar por una representación real mediante el uso de sistemas con lógica difusa [Liu et al., 2010; Cintula, Fermüller, & Noguera, 2015].

Existen dos formas para representar el conjunto de reglas difusas mediante relaciones difusas [Ciaramella et al., 2010 ; Eiter et al., 2008].

- Hacer una unión de todas las relaciones que representan a todas las reglas, o
- Hacer una unión de la salida difusa de todas las reglas.

Siendo la segunda, la forma más eficiente computacionalmente y la más empleada en los sistemas computacionales [Bragaglia, Chesani, & Ciampolini, 2010]. En la presente investigación se hará uso de las reglas difusas de tipo Mamdani para la generación de recomendaciones [E. H. Mamdani & Assilian, 1975; E. Mamdani, 1988], ya que son ampliamente aceptadas y adaptadas a la captura del conocimiento experto. Estas reglas difusas, que se define mediante un conjunto de reglas si-entonces o reglas bayesianos, se expresan de la siguiente manera:

$$(1) \quad R^m: \text{IF } u_1 \text{ is } A_1^m \text{ AND } u_2 \text{ is } A_2^m \text{ AND } \dots u_p \text{ is } A_p^m, \\ \text{THEN } v \text{ is } B^m$$

Con $m=1,2,\dots, M$ y donde A_i^m y B^m son conjuntos difusos en $U_i \subset \mathfrak{R}$ (números reales) y, $V \subset \mathfrak{R}$ respectivamente, $u = (u_1, u_2, \dots, u_n) \in U_1 \times U_2 \times \dots \times U_n$ and $v \in V$, and $x = x_1, x_2, \dots, x_n \in U$ and $y \in V$ son valores numéricos específicos de u y v , también respectivamente. Una regla de este tipo expresa una relación entre los conjuntos A y B cuya función característica sería $\mu_A \rightarrow B^2(x, y)$ y, que además representa lo que se conoce como la implicación lógica. Un ejemplo de este tipo de reglas se presentan más adelante, específicamente en dos casos de estudio para la validación del modelo conceptual multi-dominio propuesto en esta investigación.

Por otra parte, en el diseño de un sistema de lógica difusa para la generación de recomendaciones, el diseñador tiene que decidir cuál de las tres funciones de pertenencia más populares se utilizará: triangular, Gauss o trapezoidal [Negnevitsky, 2005]. En la presente tesis doctoral, se hará uso de las funciones de pertenencias triangulares y trapezoidales para la representación de los conjuntos difusos.

El Uso de las funciones triangulares y trapezoidales significa que la tasa de rendimiento será muy rápido, aunque el nivel de exactitud será menor que con cualquiera de las funciones de pertenencia [Xie, Xiong, & Church, 1998].

La función triangular se define por un límite inferior a , un límite superior b , y un valor de m , donde $a < m < b$

$$(2) \quad \mu_A(x) = \begin{cases} 0 & \text{if } x < a, \\ \frac{x-a}{m-a} & \text{if } a < x \leq m, \\ \frac{b-x}{b-m} & \text{if } m < x < b, \\ 0 & \text{if } x \geq b. \end{cases}$$

La función trapezoidal está definido por un límite inferior a , un límite superior d , un límite inferior de soporte b , y un soporte superior límite c , donde $a < b < c < d$,

$$(3) \quad \mu_A(x) = \begin{cases} 0 & \text{if } x < a \text{ or } x > d, \\ \frac{x-a}{b-a} & \text{if } a \leq x \leq b, \\ 1 & \text{if } b \leq x \leq c, \\ \frac{d-x}{d-c} & \text{if } c \leq x \leq d. \end{cases}$$

Hay dos casos especiales de una función trapezoidal, que se llaman R funciones y L-funciones. En el ex-R los parámetros $a = b = -\infty$,

$$(4) \quad \mu_A(x) = \begin{cases} 0 & \text{if } x > d, \\ \frac{d-x}{d-c} & \text{if } c \leq x \leq d, \\ 1 & \text{if } x < c. \end{cases}$$

En los posteriores los parámetros $c = d = +\infty$,

$$(5) \quad \mu_A(x) = \begin{cases} 0 & \text{if } x < a, \\ \frac{x-a}{b-a} & \text{if } a \leq x \leq b, \\ 1 & \text{if } x > b. \end{cases}$$

La notación para los conjuntos difusos se define por [Zadeh, 1965]; Sea A un conjunto difuso definido en el universo U:

$$(6) \quad A = \{(x, \mu_A(x)) / x \in U\}$$

Este conjunto difuso en el universo del discurso U se caracteriza por una función de pertenencia $\mu_A(X)$, la función toma valores en el intervalo $[0,1]$, y puede representarse como un conjunto de pares ordenados por un elemento X y su valor de pertenencia al conjunto. Una vez definidos los conjuntos difusos, las reglas de inferencia difusas pueden ser usadas para representar la relación entre estos conjuntos difusos. En este contexto, el proceso de razonamiento difuso se basará, por una parte, en hacer inferencias a partir de hechos y relaciones difusas; y por el otro, en una combinación difusa de evidencias que actualizan la exactitud de credibilidad.

Por lo tanto, los conjuntos difusos permiten formalizar expresiones lingüísticas que típicamente contienen algún grado de ambigüedad, es decir, que proveen un método para expresar matemáticamente conceptos de fácil entendimiento, tales como, alto, frío, rápido, etc. En este sentido, es posible usar los conjuntos difusos para comprender un dominio y adaptarlo al mundo real en el que vivimos.

2.5. Discusión y conclusión del estado de la cuestión

Teniendo en cuenta los distintos modelos formales, cabe señalar que existen mayores ventajas en el uso de los modelos ontológicos semánticos basados en el conocimiento. Además son usados en SR aplicados a los medios sociales, ya que dichos modelos consiguen generar un amplio conocimiento sobre una información. Además, pueden cubrir la semántica sobre un contenido específico haciendo posible de esta manera la obtención de una extensa cantidad de información aplicada a los diferentes dominios y productos existentes en la Web.

Los modelos ontológicos se aplican con ímpetu en áreas de investigación semántica, utilizando principalmente técnicas de CB y CF. Por otra parte, cabe resaltar que ninguno de los autores presenta un modelo genérico que pueda ser aplicado a distintos dominios, que permita generar contenido sobre diversos productos sin la necesidad de modificar su base de conocimiento original y, que además, pueda ser usado por un SBC. Además, los Sistemas Basados en Conocimiento son una potente herramienta de personalización en los distintos dominios, en los que la cantidad de productos disponibles desbordan la capacidad de los actores para evaluar mediante un SR cuál de ellos les puede resultar útil o interesante.

Cabe señalar que los modelos presentados ofrecen recomendaciones exitosas para los actores en cada uno de los diferentes dominios comentados, de manera que los distintos modelos generados, ayudan a dichos actores a filtrar la información de los productos existentes en el sistema, valorando sus gustos y preferencias, además de generar respuestas relevantes en el dominio. No obstante, existe una desventaja: analizando los métodos aplicados de dichos modelos, resulta que son limitados a trabajar en un dominio específico. Debido a esto, ninguno de los modelos presentados se puede aplicar a un dominio diferente a pesar de que algunos autores presentan su investigación describiendo un modelo multi-dominio, es decir, dichos modelos se limitan a trabajar en las áreas especificadas por los autores. Por lo tanto, no son modelos genéricos que puedan aplicarse a distintos dominios sin perder la originalidad de su diseño, generando ninguna posibilidad de modificar sus bases de conocimiento sin afectar su funcionalidad.

Además, ninguno de los modelos analizados presenta un nuevo modelo multi-dominio que permita representar contenido estructurado y semántico para recomendaciones mediante el filtrado de información semántica en los medios sociales.

Por otro lado, el estudio de la lógica borrosa, ofrece soluciones al planteamiento del problema en la parte de la recomendación de contenidos, ya que es una de las soluciones que se presentan para hacer que la lógica tenga grados de verdad mediante la incorporación de probabilidades. Además, la lógica borrosa se aplica a una amplia gama de problemas del mundo real para distintos dominios. Por lo tanto, mediante la transformación de componentes y conjuntos difusos es posible comprender un dominio y adaptarlo al mundo real en el que vivimos.

Asimismo, los Sistemas de Recomendación resultan de vital importancia para el filtrado de información y la valoración de múltiples contenidos almacenados en los medios sociales y la Web semántica. Además, la personalización y caracterización del contenido permiten conocer de manera eficiente las preferencias de los usuarios sobre sus gustos y la similitud de información entre los distintos productos y servicios. Asimismo, consiguen una mayor efectividad en la búsqueda de información y, por lo tanto, una mayor satisfacción en una información personalizada que los usuarios demandan.

2.6. Sumario

El propósito de este capítulo ha sido, en primer lugar, ubicar al lector al dominio de los modelos formales basados en conocimiento semántico y de contenido estructurado que son usados en los medios sociales, dominios para el que se centra la presente tesis doctoral. Como se ha podido comprobar los modelos sirven como base hacia Sistemas Basados en Conocimiento que, a partir de la extracción de información semántica de los medios sociales pueden representar un determinado conocimiento.

Por otro lado, se da a conocer la importancia de la información accesible en la Web semántica, ya que es una Web de nueva generación en la que los contenidos son más que un enorme cúmulo de información y servicios escasamente estructurados, de manera que permite utilizar el espacio de información que es aprovechable para ser procesado por máquinas, además para proporcionar y facilitar el desarrollo de los distintos modelos semánticos para la gestión de la información en diversos dominios.

Además, la información extraída de los medios sociales es fundamental, ya que influyen sobre todo tipo de actores y entornos y, poseen además, un público que es cada vez más exigente y representativo.

Por lo tanto, los resultados obtenidos por los investigadores sobre estos estudios han confirmado que los distintos modelos semánticos existentes para recomendaciones aplicados en los medios sociales, emplean soluciones eficientes y contundentes para los dominios específicos en los que han sido diseñados. Cabe destacar que, en el análisis del estado de la cuestión, ha quedado demostrado la falta de un modelo conceptual multi-dominio que permita caracterizar productos, perfiles de usuarios y los medios sociales desde diferentes paradigmas. Del mismo modo, mediante el filtrado de información semántica de los medios sociales y productos, servir de base para la operabilidad de Sistemas Basados en Conocimiento. Asimismo, considerando el contenido caracterizado y las preferencias obtenidas del comportamiento de los usuarios almacenado en el modelo conceptual propuesto, sirva para la generación de recomendaciones multi-dominio

En el siguiente Capítulo 3 de la presente investigación, se tratará de cubrir los aspectos no contemplados hasta el momento en los modelos analizados. De manera que se describe la propuesta de solución que justifica el diseño de un modelo conceptual multi-dominio para recomendaciones mediante el filtrado de información semántica en los medios sociales.

Capítulo 3

Solución propuesta

A partir de las conclusiones que se desprenden del estado de la cuestión en el capítulo anterior. En este capítulo, por un lado se exponen las razones que justifican el diseño de un nuevo modelo conceptual multi-dominio que sirva de base a Sistemas Basados en Conocimiento y, que a partir de la información extraída de los medios sociales permita la caracterización de contenido estructurado y semántico sobre productos y servicios, además la información almacenada y las preferencias de los usuarios sea útil para recomendaciones mediante el filtrado de información a partir de un SR.

Por otro lado, a partir de las hipótesis planteadas en la presente investigación y el cumplimiento de los objetivos, se define la solución propuesta a la problemática planteada al inicio de la presente tesis doctoral, en consecuencia la validación de la justificación.

3.1. Introducción

En el capítulo anterior se ha presentado una revisión del estado de la cuestión que permite ubicar al lector en el contexto donde parte la presente investigación, además descubrir la idoneidad de los modelos formales semánticos aplicados en distintos dominios y, que son usados en los medios sociales.

En la presente investigación se pretende diseñar un modelo conceptual multi-dominio para recomendaciones mediante el filtrado de información semántica en medios sociales. De manera que implica una serie de procedimientos de carácter técnico-científico que permitan la interpretación y representación del conocimiento a partir de la información extraída de las fuentes de información que se encuentran almacenadas en la Web. Así como, el diseño de un marco computacional de trabajo que integre el modelo conceptual y los componentes necesarios que, a su vez, permitan obtener los resultados planteados en el objetivo general al inicio de la presente tesis doctoral. Estas pretensiones requieren de un proceso de investigación que permita justificar las hipótesis planteadas en la Sección 1.5.

Considerando que uno de los propósitos de las hipótesis según Tamayo, es tener claras las directrices en una investigación y, que en consecuencia, resulta imperativo llevar a cabo una evaluación o validación para su comprobación [Tamayo, 2004]. Por lo tanto, llegado a este punto es necesario describir una propuesta de validación organizada que incluya el desarrollo de las distintas aplicaciones necesarias para la comprobación de las hipótesis planteadas en la presente investigación. Asimismo, la propuesta permitirá cumplir con los objetivos específicos de la investigación.

La Web almacena un extenso mundo de diferentes formas de contenidos, además de la amplia actividad de sistemas que emplean distintas herramientas y tecnologías para un constante desarrollo de nuevos modelos. En consecuencia, dichos modelos permitan la representación del conocimiento hacia los usuarios en respuestas claras y precisas sobre distintas situaciones. Actualmente, como se ha mencionado al inicio de la presente tesis doctoral y, que además es parte de la motivación para su desarrollo, no existe un

modelo que pueda ser aplicable a distintos dominios sin modificar sus elementos formales o los conceptos principales de su base original. Que además, sirva de base a un SBC y un SR pueda usar el contenido caracterizado para generar recomendaciones personalizadas en múltiples dominios.

Actualmente los motores de búsqueda en la Web se enfrentan a problemáticas que dificultan mucho la recuperación de información en la enorme base de datos que almacena la Web [Kherfi, Ziou, & Bernardi, 2004]. Por lo tanto, se busca resolver esos problemas anotando conceptualmente el contenido. Sin duda, esto mejoraría por mucho la experiencia del usuario. Parte fundamental en la semántica es la anotación mediante la creación de ontologías, ya que pueden ser entendidas tanto por usuarios como por otros sistemas de cómputo. Las ontologías en la Web incluyen la extracción, generación de contenido. En la actualidad, esto es posible de manera manual o semi-automática [Vallet, Fernández, & Castells, 2005]. Por otro lado, para la creación de contenido se necesita el uso de tecnologías que permitan la representación del conocimiento sobre palabras, datos o conceptos en lenguajes estándares que puedan ser formalizados en un área específica de conocimiento.

La caracterización de contenido en la Web es un área de aplicación y explotación para las tecnologías y líneas de investigación que emergen constantemente [Baeza-Yates & Ruiz-del-Solar, 2004]. Gracias a la interoperabilidad que existe en la Web, permite a los usuarios interactuar y colaborar entre sí para generar un contenido. Por lo tanto, el contenido es caracterizado para ofrecer, en un futuro, resultados eficientes hacia los usuarios a través de herramientas y modelos que interaccionen entre ellos. Asimismo, la gestión de contenidos sobre distintos productos en la Web que siempre requiere de una evaluación para establecer una solución a un contexto dado.

Por lo tanto, la propuesta para el diseño de un modelo conceptual multi-dominio se describe en la siguiente sección que además, permite comentar la solución a la problemática planteada al inicio de la investigación.

3.2. Modelo propuesto

La gran variedad de información almacenada en la Web permite a las tecnologías semánticas y técnicas de aprendizaje, desarrollar un filtrado de información semántica que posibilita la convergencia entre los usuarios y los medios sociales a través de la creación de diversos contenidos [Cromity, 2012]. La problemática asociada al significado, sentido e interpretación de palabras, expresiones, símbolos o una representación formal, resultan un problema lógico de significación. Por lo tanto, la generación de reglas que permitan asegurar la organización y codificación del contenido, es de vital importancia para la gran diversidad de datos heterogéneos que se encuentran almacenados en la Web. Ya que dichos datos deben ser procesados de forma rápida, transparente y eficaz. Los modelos conceptuales son usados por SBC para representar, organizar y codificar la información y, a su vez, permiten el análisis de los datos mediante técnicas de filtrado de información sobre diversos dominios de aplicación. Por consiguiente, son capaces de generar y gestionar un contenido que se basa en reglas o conocimiento, las cuales, sin embargo, y en algunas ocasiones, no siempre son correctas y válidas. No obstante, cabe mencionar que la representación del conocimiento a través de reglas difusas permite una eficaz representación lógica en los SBC [Colomo-Palacios, 2012 ; García-Crespo et al., 2012].

Por otro lado, los usuarios demandan cada vez y con mayor frecuencia un conocimiento transparente sobre la información relativa a los contenidos en la Web. Por lo tanto, llegado a este punto es necesario recordar el objetivo principal de la presente investigación que respalda la definición del modelo propuesto.

El objetivo principal de la investigación es la creación de un nuevo modelo conceptual genérico multi-dominio que servirá como base hacia un Sistema Basado en Conocimiento para representar, a partir de la extracción de información semántica de los medios sociales, información sobre los distintos productos que satisfacen las necesidades de un usuario, a su vez, modelará y gestionará el conocimiento de diferentes perfiles de usuarios, productos y medios sociales caracterizados en múltiples dominios. Además, el modelo permitirá, a partir de un SR, generar recomendaciones personalizadas para los usuarios. Esta representación se enfoca en la información almacenada en las fuentes de

información de los medios sociales, la caracterización de productos y la recomendación personalizada, además de ser trasladado a diversos dominios.

Por lo tanto, el modelo debe de representar el conocimiento a partir de la extracción de la información almacenada en las plataformas de comunicación social en la Web mediante la utilización de las TS, técnicas de aprendizaje y procesos de filtrado o recuperación de la información. Un Sistema Basado en Conocimiento podrá hacer uso del modelo como una base para utilizarlo en procesos de decisión en dominios específicos.

El conocimiento que se encuentra en las distintas fuentes de información dará el soporte hacia un experto en el dominio que proporcione una información precisa que pretenda ser modelada. La manera de adquirir el conocimiento será mediante una herramienta de modelado que permite extraer la información y representarla en una base de conocimiento. Como se ha comentado anteriormente en el capítulo del estado de la cuestión, la Web semántica proporciona herramientas y tecnologías adecuadas para el estudio de los distintos modelos existentes basados en conocimiento aplicados a diversos dominios. Por lo tanto, las ontologías forman la base para el diseño del modelo conceptual multi-dominio propuesto en esta investigación.

Por otro lado, la información será unida en un sistema que es soportado en un marco computacional de trabajo. Además, el modelo conceptual multi-domino debe tener la capacidad de representar las características de diversos productos, perfiles y redes sociales.

El modelo conceptual propuesto se basa en:

1. La representación del conocimiento extraído de los medios sociales en la Web.
2. La capacidad semántica y de soporte para contenido estructurado con características multi-dominio.
3. La creación de una base de conocimiento para la caracterización de diversos contenidos a ser procesados por un Sistema Basado en Conocimiento.

4. La representación de las recomendaciones en entornos multi-dominios que son soportadas por la operabilidad de determinados expertos en el dominio, pudiendo ser personas u otros sistemas que intercambien la información relevante.
5. Un motor de razonamiento difuso capaz de evaluar las preferencias de los usuarios a partir de reglas establecidas que permitan una representación lógica en el contenido estructurado semántico del sistema.
6. Una plataforma de trabajo que integra los distintos componentes para una solución sobre los contenidos en la Web.

3.3. Sumario

Como se ha mencionado en esta sección, se han expuesto las razones que justifican el desarrollo de la presente tesis doctoral, además de comentar una vez más el planteamiento del problema presentado al inicio de la tesis doctoral y la solución a partir de las hipótesis planteadas que respaldan las directrices a seguir y, que además permitan alcanzar la validación de la presente investigación. Como resultado se introduce la concepción de un nuevo modelo conceptual multi-dominio para la representación del conocimiento.

En la siguiente sección que corresponde al Capítulo 4, se describe el modelo conceptual multi-dominio para la representación del conocimiento, que es el tema central de esta tesis doctoral. Además, las consideraciones para el desarrollo de una arquitectura basada en el modelo propuesto y, comentar los componentes a ser usados durante el desarrollo.

Capítulo 4

Modelo conceptual multi-dominio

Un modelo conceptual es una abstracción, una vista simplificada del mundo que se desea representar para algún propósito. Cada base de conocimiento, agente o sistema basado en conocimiento establecen relaciones entre conceptos definidos en ciertos niveles de abstracción que se asocia con alguna conceptualización explícita o implícitamente de un caso particular.

En este capítulo se diseña un modelo conceptual para recomendaciones mediante el filtrado de información semántica en redes sociales. En primer lugar, se realiza una definición formal del modelo y los conceptos abstractos que lo definen, En segundo lugar, el modelo es representado en distintas vistas parciales. En tercer lugar, la definición de una ontología que describe las relaciones entre los distintos conceptos principales del modelo.

En consecuencia, a partir del propuesto modelo conceptual es posible representar un conocimiento de diferentes dominios, además permite la caracterización de la información sobre los distintos productos y servicios que satisfacen las necesidades de un usuario.

4.1. Consideraciones previas

Antes de abordar la definición formal del modelo conceptual propuesto y sus conceptos principales que lo definen, es necesario considerar las definiciones abstractas sobre el término de modelo conceptual y el punto hacia donde se pretende abordar la solución del problema. Además, se van a plantear las consideraciones generales que sirvan para la comprensión de los distintos conceptos que van a ser utilizados en el modelo multi-dominio propuesto. Posteriormente, se realizará la conceptualización del modelo multi-dominio definiendo los conceptos principales, propiedades y relaciones que existen entre ellos. Para la conceptualización del modelo se ha decidido emplear uno de los lenguajes de ontologías OWL, ya que permite describir las características del modelo y el conocimiento sobre un dominio con mayor expresividad. Asimismo, dicha ontología que define el modelo se representará en forma de una base de datos relacional.

Por otro lado, en esta investigación se plantea la representación del conocimiento a partir de la información extraída de los medios sociales. En consecuencia, un experto en el dominio facilitará el conocimiento sobre ciertas características de los productos de un dominio. Uno de los objetivos planteados al inicio de la investigación es la representación del conocimiento para múltiples dominios. De tal forma que, en la sección de validación se han planteado dos casos de estudio que permitan llevar a cabo por un lado, la validación del modelo propuesto y por otro lado, la demostración de las hipótesis planteadas en la investigación, para ello además, es necesario plantear el diseño de un marco computacional de trabajo basado en el modelo conceptual que permita conocer la validez del modelo propuesto y, al mismo tiempo, permita ofrecer resultados concretos en la información para múltiples dominios. Por lo tanto, el uso de las TS, técnicas de aprendizaje y procesos de filtrado de información permitirán la definición y diseño de un nuevo modelo conceptual genérico propuesto en esta investigación. En consecuencia, es importante mencionar las distintas definiciones que hacen referencia a los distintos modelos conceptuales existentes en la literatura.

El término modelo conceptual, se utiliza en varios dominios, incluyendo la ingeniería del conocimiento y la filosofía de la ciencia que investiga el conocimiento

científico y la práctica científica, por lo que se constituyen en una de las disciplinas más complejas y abstractas existentes actualmente. Los modelos conceptuales tal y como se entiende en la ingeniería, son un tipo de modelos con una descripción abstracta. No obstante, llevan utilizándose muchos años en el área de la ingeniería y la ciencia para su uso con distintos propósitos [Musen, 1989].

Un modelo conceptual, es una representación abstracta de algo generalizado en casos particulares [Borah, 2002]. Además, Pace menciona que un modelo conceptual implica la construcción de representaciones del conocimiento humano [Pace, 2000]. Complementariamente, Robinson proporciona una definición más amplia para un modelo conceptual, en ella indica que se trata de una descripción Non-Software específica sobre un modelo a desarrollar, describiendo sus objetivos, sus entradas y salidas, el contenido, los supuestos y las simplificaciones del modelo [Robinson, 2006].

Adicionalmente, Robinson menciona que un Modelo Conceptual tiene las siguientes propiedades [Robinson, 2006]:

- La actividad del modelado conceptual es iterativa y repetitiva a través de todo el ciclo de su desarrollo.
- El modelo conceptual es una representación simplificada y abstracta de un sistema real.
- El modelo conceptual es independiente del código del modelo o del Software.
- Las perspectivas tanto de los usuarios como de los desarrolladores, son tomadas en consideración.

De manera que el diseño de un modelo conceptual tiene grandes ventajas prácticas. Por ejemplo, simplifica lo que se describe de diversas formas permitiendo a su creador centrarse en los aspectos que considera importantes, de manera que un modelo

conceptual se transforma en un instrumento de reducción de la complejidad. Además, independientemente de los tipos de modelos conceptuales que existen o de sus propiedades formales, un modelo conceptual tienen un propósito fundamental, ayudar a favorecer la comprensión sobre un tema y proyectar o anticipar la estructura o funcionamiento del mismo [Thompson, Higgins & Howell, 1991].

Antes de describir los conceptos que definen el modelo de la presente tesis doctoral, es importante considerar el objetivo y los puntos de vista desde el que se quiere abordar la solución al problema. De manera que, el modelo conceptual multi-dominio propuesto para esta investigación, tiene como objetivo representar la información sobre los distintos productos que satisfacen las necesidades de un usuario, a su vez, modelar y gestionar el conocimiento de diferentes perfiles de usuarios, productos y medios sociales caracterizados en múltiples dominios. Además, el modelo conceptual propuesto debe servir como base hacia un Sistema Basado en Conocimiento.

Como se ha mencionado en los estudios descritos en la Sección 2.2 y, abordando la problemática sobre los modelos formales semánticos Basados en conocimiento. El resultado que se aprecia en los modelos presentados con anterioridad, resulta que presentan una desventaja, son modelos que no pueden ser trasladados a un dominio distinto para el que fueron diseñados originalmente, limitándose por tanto, a su uso en áreas específicas y restringidas para los SR. Para ello, en esta investigación se plantea un modelo multi-dominio capaz de representar el conocimiento desde distintos dominios. Para ello, es necesario tener en cuenta las consideraciones que permitan dar una solución al planteamiento del problema en esta investigación.

A partir de estas consideraciones previas, el diseño del modelo se llevará a cabo a partir de la definición de los conceptos de representación del conocimiento, esto conceptos se definirán en distintas vistas parciales de abstracción, la primera será la definición de un modelo abstracto, la segunda será la definición denominada modelo conceptual multi-dominio. Finalmente, una vista parcial del modelo conceptual la cual denominaremos modelo formal que implica la descripción de los principales conceptos y relaciones del modelo conceptual multi-dominio.

Por lo tanto, a continuación se describirá el modelo conceptual en una vista que denominamos modelo abstracto para la representación del conocimiento. Dicha vista, permite visualizar el proceso de extracción y la clasificación del contenido a partir de las distintas fuentes de información almacenadas en la Web.

4.2. Modelo abstracto

El modelo conceptual multi-dominio propuesto, se basa en una representación del conocimiento a través de conceptos presentados en distintas vistas parciales. En este sentido, cada uno de los conceptos es descrito a detalle en la Sección 4.4. Además, el conjunto de conceptos que forman el modelo propuesto, sus propiedades y sus relaciones son representados en una ontología para la representación del conocimiento, la cual es descrita en siguientes secciones.

La Figura 4:1 muestra la arquitectura de alto nivel para la representación del conocimiento del modelo abstracto y la información extraída de las distintas fuentes de información en la Web.

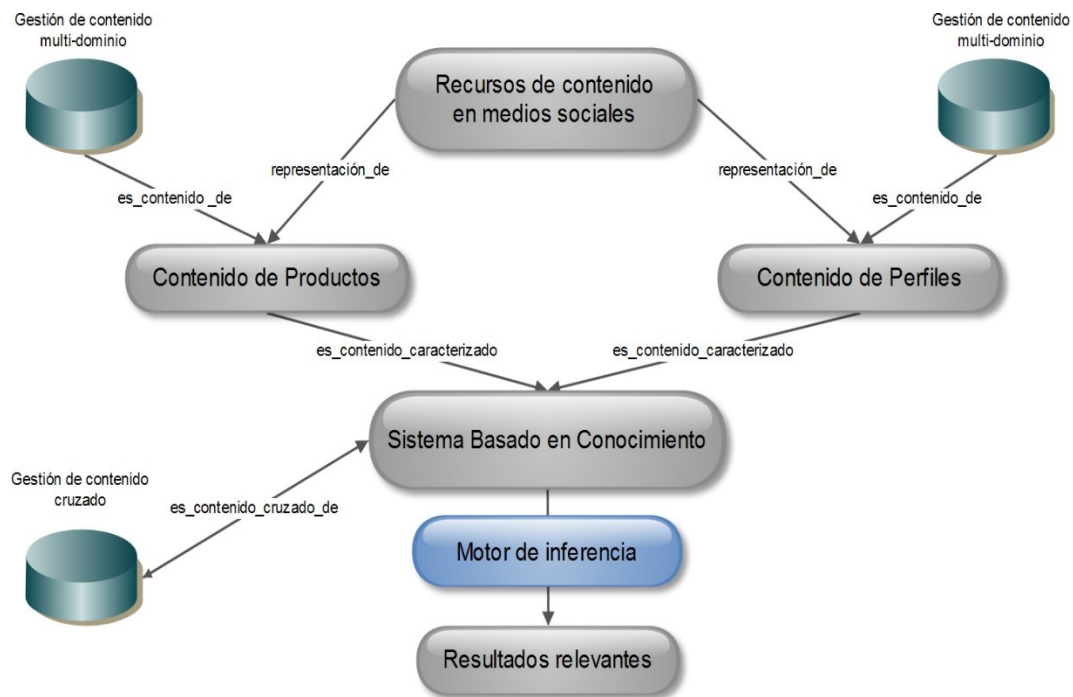


Figura 4:1 Modelo abstracto para la representación del conocimiento.

El modelo abstracto representa la administración de contenido extraído a partir de las distintas fuentes de información en los medios sociales disponibles en la Web. La información permite realizar la caracterización de un contenido, para este caso, sobre productos y perfiles. Por otro lado, un sistema basado en el conocimiento integrado en un marco de trabajo y, que además, forma parte del filtrado de información para la generación de recomendaciones en relación con el contenido caracterizado, gestiona el contenido semántico del sistema. De manera que, mediante la utilización de un razonador proporciona los resultados relevantes a ser usados en diferentes dominios.

4.3. Modelo conceptual multi-dominio

En la sección anterior se han tenido en cuenta las consideraciones necesarias para diseñar un modelo abstracto, el cual conlleva a la definición de una vista parcial denominada modelo conceptual multi-dominio.

El modelo conceptual multi-dominio representa la definición concreta de conceptos que permiten gestionar una representación del conocimiento para diferentes dominios, los conceptos son descritos posteriormente y, definidos a detalle en la Sección 4.4. Además, a raíz de la definición del modelo conceptual se presenta una vista parcial a la que denominaremos modelo formal, de la cual, se desprenden distintas vistas del modelo multi-dominio propuesto en esta investigación para la representación del conocimiento.

A partir de estas definiciones es necesario considerar los supuestos y restricciones que establecen un marco de trabajo como base hacia un Sistema Basado en Conocimiento, que a su vez, permita representar la gestión del conocimiento de diferentes perfiles de usuarios, productos y medios sociales caracterizados. Asimismo, a partir de un SR integrado en el marco de trabajo, permita la generación de recomendaciones a partir de la información almacenada en las plataformas de comunicación social en la Web. El proceso es realizado mediante la utilización de las TS, técnicas de aprendizaje y los procesos de filtrado o recuperación de la información semántica.

Por lo tanto, el modelo conceptual multi-dominio propuesto para esta investigación, permite modelar conceptos relacionados a una red social, a un producto y a un perfil de usuario para la caracterización del contenido, para su descripción se ha utilizado el lenguaje ontológico. Una recomendación de la W3C, que es ampliamente utilizado en el dominio de la Web semántica y las ontologías. La Figura 4:2 muestra el diseño de un nuevo modelo conceptual multi-dominio para recomendaciones mediante el filtrado de información semántica en medios sociales propuesto en la presente investigación.

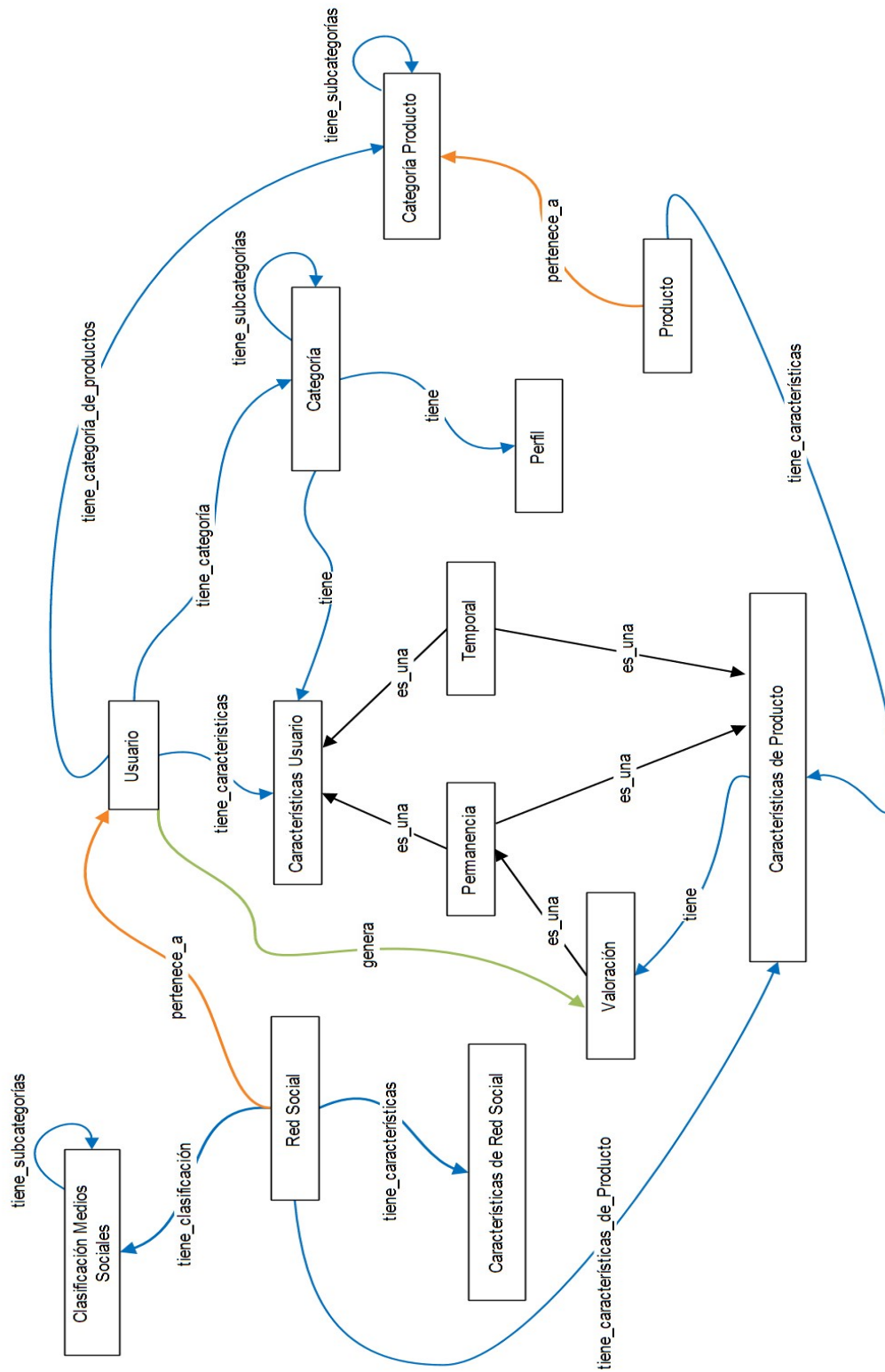


Figura 4:2 Modelo conceptual multi-dominio.

En la Tabla 4.1 se desglosan cada uno de los conceptos del modelo conceptual multi-dominio y se describe un ejemplo para una mejor comprensión sobre su utilización de cada uno de los conceptos definidos.

Tabla 4:1:
Descripciones y ejemplos para los conceptos definidos en el modelo conceptual multi-dominio

Concepto	Definición	Ejemplo
Usuario	Actores activos y pasivos que interactúan en la web para editar, publicar y compartir información.	User@tweetero nombre ficticio
Red Social	Plataforma de comunicación en la Web. Su estructura permite la interacción dinámica entre los usuarios.	Twitter
Producto	Un conjunto de atributos tangibles e intangibles que se ofrecen desde una línea en el mercado y, además cumplen con una necesidad.	Marcas / Servicios
Perfil	Un conjunto de características asociadas con un elemento o categoría.	Tímido, emocional, racional
Clasificación de medios sociales	Tipos de plataformas de comunicación donde el contenido es creado por el uso de tecnologías de la Web.	Redes sociales, blogs, microblogs, videojuegos, Fotos, Wikis, podcasts, foto y vídeo.
Características de usuario	Los usuarios pueden distinguirse de los demás en función de diversas características	Educación / Experiencia
Características de producto	Los productos pueden ser distinguidos de los demás en función de diversas características	Diseño / Motor / Seguridad
Características de red social	Red Social se puede distinguir de los demás en función de diversas características	Microblogging, mensajes de texto
Permanencia	Es un concepto que nace de las características del usuario, por lo que puede o no puede ser permanente	Introvertido / Corporativo
Temporal	Es un concepto que nace de las características del usuario y las características del producto, por lo que puede o no puede ser temporal	Ubicación geográfica / Fecha de Caducidad
Categoría	Representa el contenido de una clase asociada a un perfil caracterizado producto y caracterizado, el	Ecologista / Seguridad

	contenido se infiere de la información almacenada en las redes sociales	
Valoración	Representa al contenido de una clase asociada con las preferencias de un usuario y el producto característico, el concepto permite las calificaciones de un producto en particular a través de usuario	Gustos y preferencias
Categoría de productos	Es un concepto que nace del contenido asociado en la Web, hay sub-categorías que pueden almacenar un contenido asociado a esas categorías descritas originalmente	Negocio / Consumo

Las relaciones entre cada uno de los conceptos del modelo conceptual propuesto, permiten representar, caracterizar y gestionar el conocimiento para diferentes perfiles, productos y medios sociales en distintos dominios.

En este sentido, a continuación se describirá un modelo formal, el cual representa los conceptos principales del modelo conceptual multi-dominio que permite ser trasladado a distintos dominios sin cambiar su base de conocimiento.

4.4. Modelo formal

El modelo formal representa una vista parcial abstracta sobre los conceptos principales que comprenden al modelo conceptual multi-dominio, además el modelo formal desglosa distintas vistas parciales para la representación del conocimiento. Antes de presentar las distintas vistas parciales es importante describir cada uno de los conceptos propuestos para la definición del modelo conceptual multi-dominio.

Por lo tanto, las siguientes secciones describen detalladamente los conceptos principales introducidos en el modelo conceptual multi-dominio.

4.4.1. Producto

Un producto es una opción elegible, viable y repetible que la oferta pone a disposición de la demanda, para satisfacer una necesidad o atender un deseo a través de su uso o consumo [Kotler & Gertner, 2002]. El concepto <<Producto>> representa una clase dentro del modelo para la clasificación de un conjunto de atributos tangibles e intangibles que se ofrecen desde una línea en el mercado y, que además, cumplen con una necesidad. En los SR los productos se describen por una lista de características, normalmente representado por un vector en el que cada posición del vector indica una característica y su valor, el peso o importancia que tiene para el usuario dicha característica [Mak et al., 2003]. La Figura 4:3, representa un fragmento de la taxonomía semántica del modelo, en el fragmento se muestra la clase <<Categoría de producto>> y la clase <<Producto>> que pertenece a una categoría de productos. Este concepto denominado producto, permite modelar y representar el conocimiento en un dominio, además de tener características que permiten realizar una caracterización sobre los distintos productos disponibles en la Web.

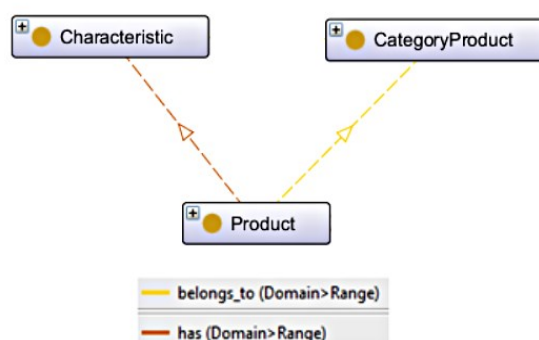


Figura 4:3 Fragmento de la taxonomía para un criterio denominado <<Producto>>.

En este sentido, la anotación semántica de un producto permite encontrar los atributos que hacen explícito lo que el usuario percibe y utiliza para juzgar cada producto, a partir de formas, deseos, emociones, colores, recuerdos, etc. Además, Vergara &

Mondragón presenta una taxonomía completa de un producto que permite visualizar las características que representa un producto [Vergara & Mondragón, 2014]. Por lo tanto, el concepto permite representar cualquier producto o marca almacenado en la Web y, al mismo tiempo anotarlo semánticamente en la ontología del sistema.

4.4.2. Categoría

Una categoría en un lenguaje coloquial, se entiende como un nivel de importancia de cualquier cosa respecto a las demás. Por otra parte, mediante las categorías se demuestra una clasificación jerárquica sobre productos y servicios, además permiten clasificar o agrupar productos similares [Noy, Sintek, & Decker, 2001]. Por lo tanto, el concepto <<Categoría de usuario>>, representa el contenido de una clase asociada a un perfil caracterizado, denominado <<Perfil>>. El contenido se infiere de la información almacenada en las redes sociales y el perfil se crea a partir de los criterios que el experto establece en el dominio. En la Figura 4:4, se muestra un fragmento de la taxonomía que especifica al concepto de categoría.

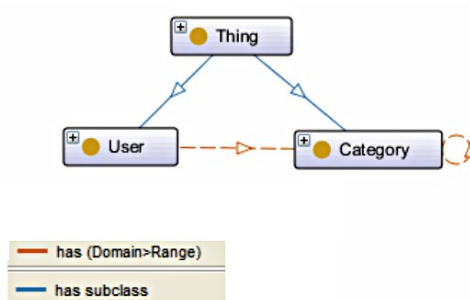


Figura 4:4 Fragmento de la taxonomía para el concepto <<Categoría>>.

4.4.3. Perfil

Un perfil es un conjunto de características asociadas con un usuario que, asimismo, representa con precisión los intereses de los usuarios [Trajkova & Gauch, 2004]. Para este caso, un perfil se encuentra asociado con el concepto denominado <<Categoría>>.

El concepto permite filtrar la información de un perfil que sirve para construir un nuevo contenido a partir de la información almacenada en la memoria, a su vez, permite modelar los gustos y preferencias de los usuarios a partir de la información caracterizada y anotada semánticamente en la base de conocimiento.

Por lo tanto, los datos permiten gestionar y establecer un contenido caracterizado en la creación de un nuevo perfil que, a su vez, sirva para la generación de recomendaciones personalizadas. En la Figura 4:5 se muestra un fragmento de la taxonomía usada para el concepto de <<Perfil>>. La clase permite definir distintos tipos de perfiles a partir del experto en esta investigación.

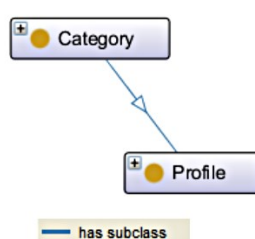


Figura 4:5 Fragmento de la taxonomía para un concepto denominado <<Perfil>>.

4.4.4. Categoría de productos

A diferencia del concepto <<Categoría>> antes descrito, este concepto denominado <<Categoría de productos>> almacena los diversos dominios a ser

modelados, es decir, que el concepto es genérico multi-dominio ya que puede representar el conocimiento de cualquier dominio de contenido estructurado y semántico. Del mismo modo, el concepto está asociado al concepto denominado <<producto>>. Un fragmento de la taxonomía sobre este concepto se muestra en la Figura 4:6.

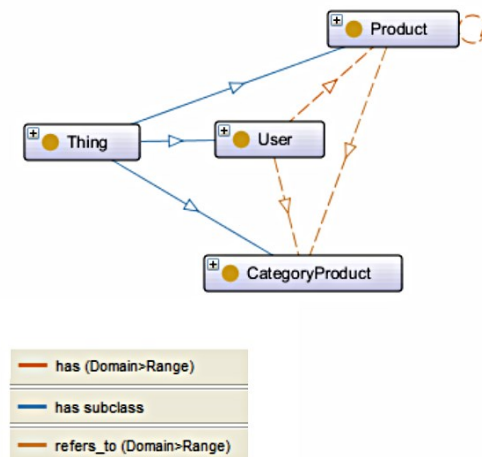


Figura 4:6: Fragmento de la taxonomía para un concepto denominado <<Categoría de productos>>.

4.4.5. Características

Es un concepto que puede ser distinguido a partir de los demás en función de diversas características. Es decir, los productos, los usuarios y las redes sociales contienen atributos que son características específicas. Además, a través del caracterizado de un producto o un servicio se obtienen distintos beneficios de interés sobre una necesidad o un deseo esencial del usuario hacia un servicio, producto o contenido [Hassanein & Head, 2014]. En la Figura 4:7, se muestra un fragmento de la taxonomía que representa las características de un producto, un usuario y una red social definida para el desarrollo de esta investigación.

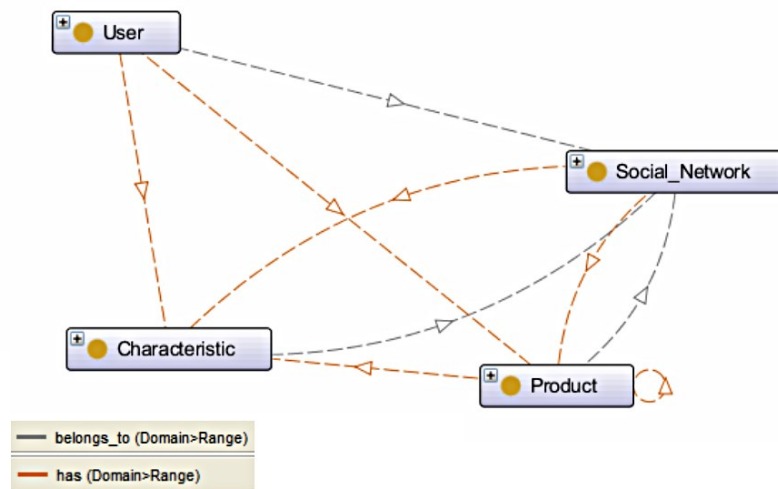


Figura 4:7 Fragmento de la taxonomía para un concepto denominado características.

Por lo tanto, la representación semántica para la caracterización de productos, servicios en la Web, usuarios y redes sociales, es de gran importancia para el modelado que pueda cubrir las necesidades de los distintos usuarios en los extensos escenarios de información [Hepp, 2008].

4.4.6. Permanencia

Es un concepto que nace a partir de las características específicas de un usuario hacia un producto. Es decir, una característica que se mantiene constante y que no cambia con el tiempo, además por lo que puede o no puede ser permanente. En la Figura 4:8 se muestra un fragmento de la taxonomía que representa la característica para el concepto denominado <<Permanencia>>.

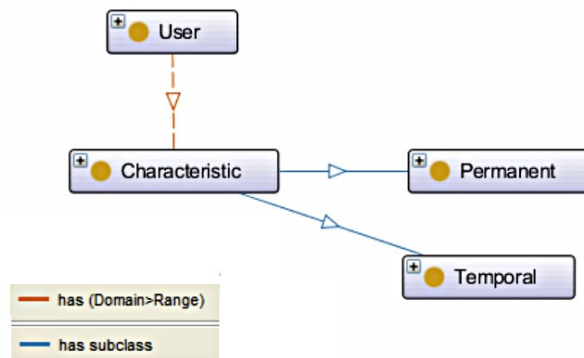


Figura 4:8 Fragmento de la taxonomía para un concepto de caracterizado <<Permanencia>>.

4.4.7. Temporal

Al igual que el anterior concepto descrito <<Permanencia>>, es un concepto que nace a partir de las características específicas de un usuario y las características de un producto. Dicha característica puede cambiar con el tiempo, por ejemplo una localización geográfica. Asimismo, el concepto <<Temporal>> puede o no puede ser temporal. En la Figura 4:9 se muestra un fragmento de la taxonomía que representa la característica para el concepto denominado <<Temporal>>.

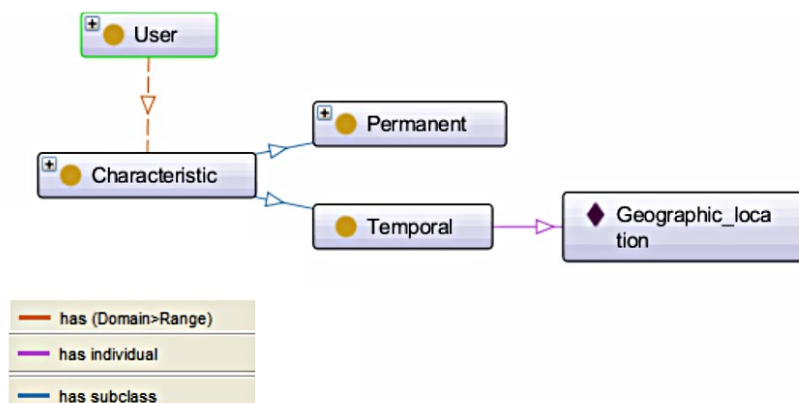


Figura 4:9 Fragmento de la taxonomía para un concepto de caracterizado <<Temporal>>

4.4.8. Valoración

La valoración o rating, es un término usado para hacer referencia a una calificación sobre un cierto elemento, es un suministro de información, su labor consiste en emitir una valoración o calificación a través de expertos sobre las opiniones de usuarios sobre una calidad crediticia. En este caso, el concepto <<Valoración>> representa al contenido de una clase asociada con las preferencias de un usuario hacia un producto caracterizado, el concepto permite valorar las calificaciones sobre un tema o características en particular, a través de la generación del usuario. Es decir, el concepto mantiene una valoración que genera el usuario. En la Figura 4:10, se muestra un fragmento de la taxonomía que representa la valoración del experto hacia un producto a partir del caracterizado, además de las características de interés para el usuario definidas en el modelo conceptual.

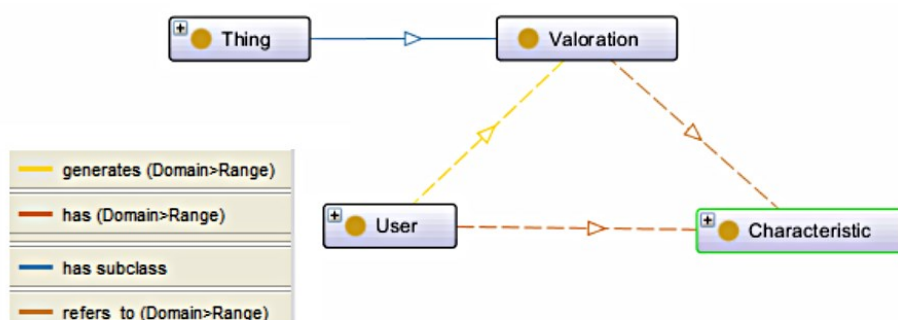


Figura 4:10 Fragmento de la taxonomía para el concepto de << Valoración>>.

4.4.9. Clasificación y características de red social

En el modelo conceptual propuesto se define un concepto denominado <<Red social>> este concepto nace a partir de las distintas plataformas de comunicación social

existentes en la Web. El concepto almacena la red social a ser modelada, además existe la necesidad de caracterizar y clasificar el concepto.

Por lo tanto, se denominan los conceptos de <<Características de red social>> y <<Clasificación de red social>> que representan a conceptos relacionado con la red social. De manera que, la información permite modelar los contenidos de una red social y se distingue a partir de sus características y su clasificación. En la Figura 4:11, se muestra un fragmento de la taxonomía que representa la clasificación y características de una red social definida en el modelo conceptual multi-dominio propuesto para esta investigación.

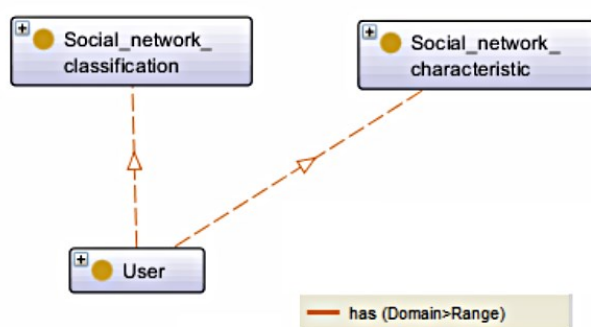


Figura 4:11 Fragmento de la taxonomía para la clasificación y características de una red social

En las secciones anteriores se han identificado los conceptos principales del modelo propuesto en esta tesis doctoral. Además, se han descrito su contexto al que se refieren cada uno de ellos dentro del modelo formal. En la siguiente sección se describen las relaciones entre los principales conceptos que comprenden al modelo formal propuesto en esta investigación.

4.5. Vistas parciales del modelo conceptual multi-dominio.

4.5.1. Relaciones entre los conceptos principales del modelo conceptual multi-dominio

A partir de la definición de los conceptos descritos en la Sección 4.3, se desprenden las relaciones que existen entre cada uno de los conceptos que comprenden una vista parcial abstracta del modelo conceptual multi-dominio. Por lo tanto, es necesario describir las posibles relaciones que existen entre ellos, la semántica y la forma que se representan para cada uno de ellos.

A continuación se explican a detalle las relaciones entre cada uno de los conceptos que se han definido en el modelo formal. En la Figura 4:12, se muestra una vista parcial abstracta sobre los conceptos principales del modelo, el cual se ha denominado, modelo formal.

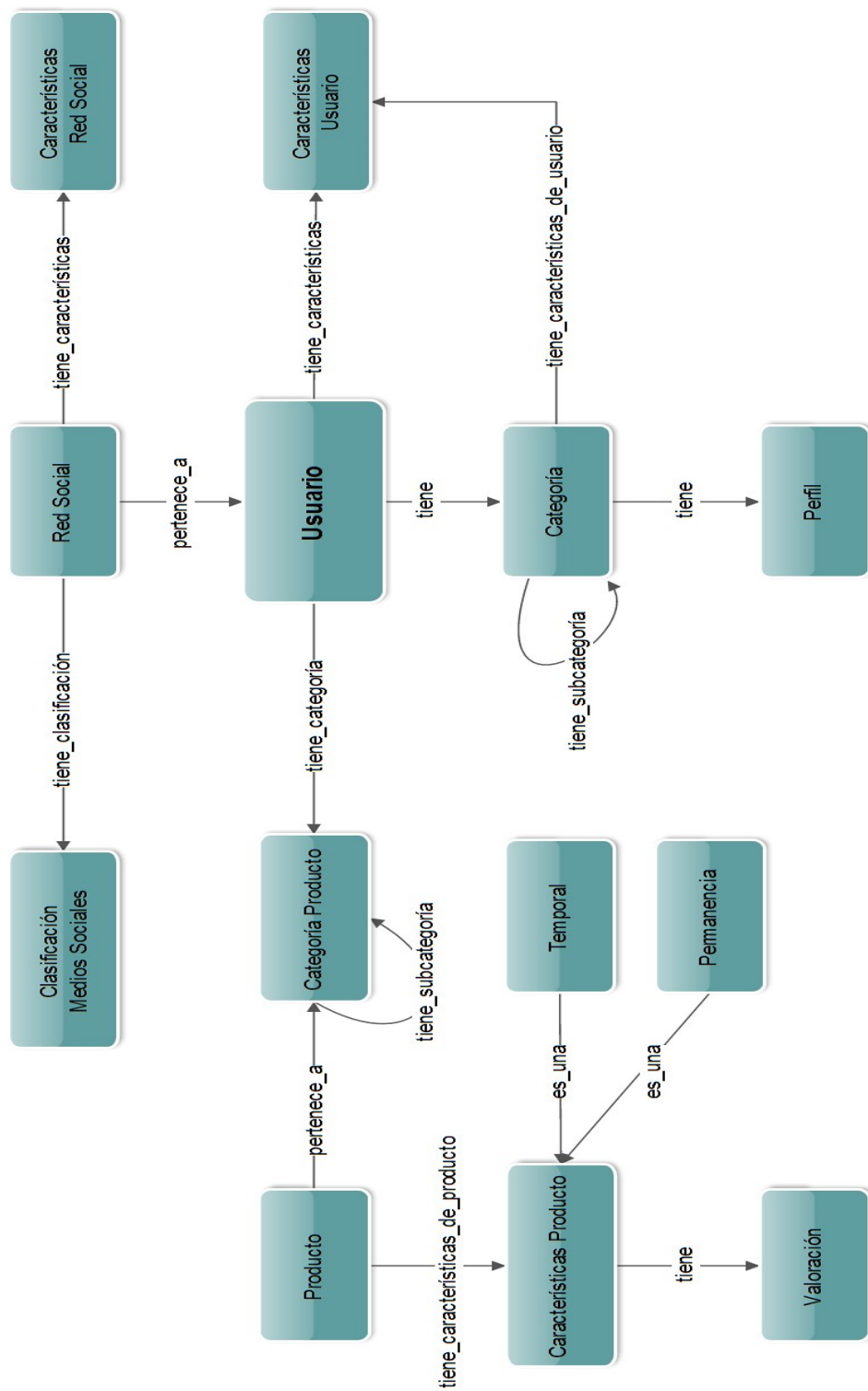


Figura 4:12 Vista parcial abstracta denominada modelo formal.

Para poder modelar el conocimiento de perfiles de usuario y productos caracterizados a partir de la información extraída de los medios sociales, es necesario establecer relaciones entre los principales conceptos definidos en el propuesto modelo. Además, en el modelo formal el usuario tiene conceptos establecidos que permiten definir las distintas relaciones entre ellos. Por lo tanto, a continuación se describen de manera general cada uno de los conceptos principales del modelo que mantienen una relación y, que además permiten producir un contenido semántico para ser modelado. Los conceptos denominados <<Producto>> y <<Categoría>>, ambos mantienen una relación con el concepto llamado <<Usuario>>. En consecuencia, los conceptos denominados <<Red social>> y <<Perfil>>, mantienen una relación con el concepto definido como <<Categoría>>. Asimismo, una red social se encuentra relacionada con un concepto definido como <<Clasificación de medios sociales>>.

Con las relaciones descritas entre los conceptos generales del modelo formal, es posible realizar una caracterización a partir de la información extraída de los medios sociales y, que a su vez, dicha información se encuentra almacenada en la Web semántica. Por lo tanto, a continuación se desglosan de manera detallada las posibles relaciones entre los principales conceptos del modelo formal.

En la Figura 4:13, se describe la relación entre los conceptos de <<Usuario>> y <<Categoría>>. Además, los conceptos de <<Producto>> y <<Red social>>, se encuentran asociados al concepto <<Categoría>>. La relación permite definir los conceptos de producto y red social como un tipo de categoría que se encuentra relacionada con un usuario. En consecuencia, dichos conceptos mantienen características específicas entre ambos, ya sean características sobre un producto o una red social. Por lo tanto se define un concepto denominado características, ya sea de productos o de una red social, el concepto permite generar una relación para la caracterización de contenido entre ambos conceptos descritos anteriormente.

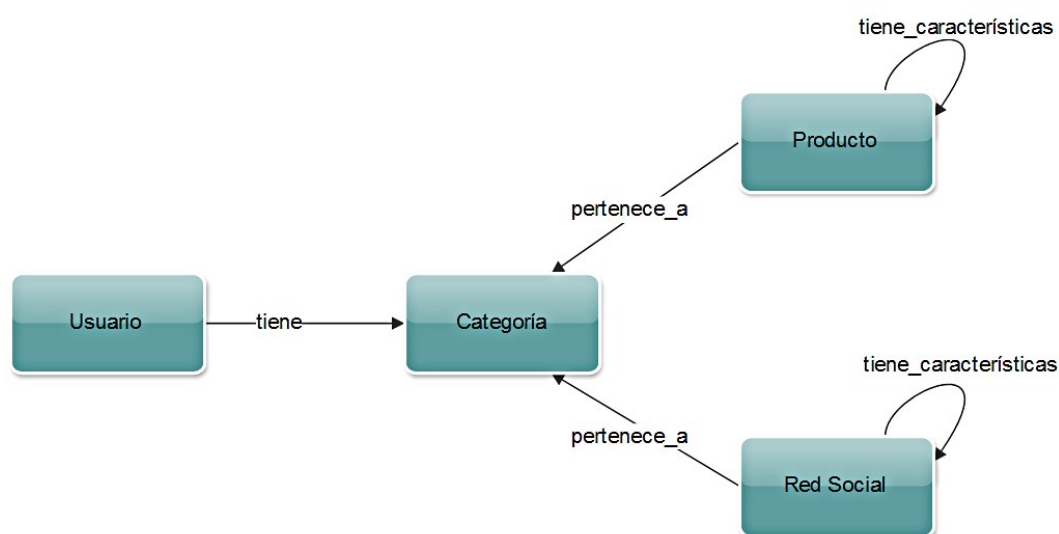


Figura 4:13 Relaciones para el concepto definido como <<Categoría>>.

En este sentido, se definen las relaciones para un concepto denominado <<Características>>. Las relaciones permiten identificar un caracterizado de los conceptos de <<Producto>> y <<Red social>>. Las características que contienen ambos conceptos pueden o no pueden estar asociadas entre sí. Además, los conceptos mantienen una relación con el concepto denominado <<Usuario>>. En consecuencia, un producto mantiene una relación con el concepto definido como <<Categoría de productos>>. Cabe resaltar que este concepto es genérico, lo que permite al modelo conceptual tener una amplia información adaptable a cualquier dominio de contenido estructurado y de similitud de contenidos almacenados en los medios sociales y que son ofertados entre los distintos usuarios. El concepto permite alcanzar uno de los objetivos específicos planteados al inicio de la tesis doctoral. Por lo tanto, el concepto permite identificar cualquier tipo de producto o servicio almacenado en la Web.

Del mismo modo, se han definido las relaciones para el concepto <<Categoría de productos>>, ya que el concepto permite por un lado, clasificar las características de un producto, por otro lado las características de una red social que pueden o no pueden mantener las mismas características, que además estén asociadas con los intereses de un usuario. Por lo tanto, la Figura 4:14, muestra las relaciones de los conceptos descritos

y la creación de un concepto definido como <<Características del producto>>, estos conceptos sirven para la gestión del conocimiento en el modelo formal.

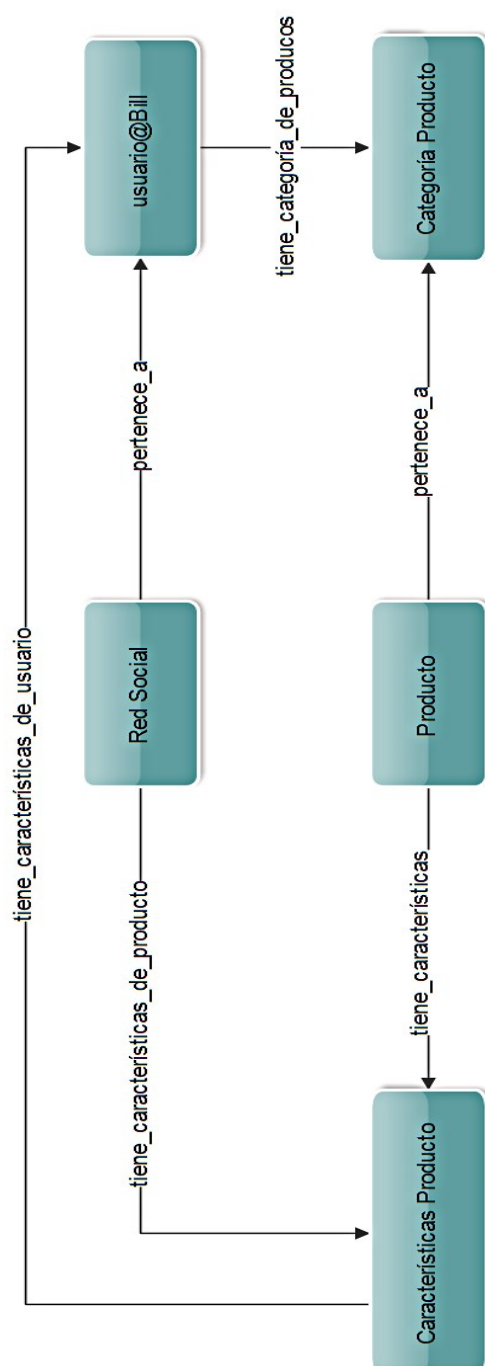


Figura 4:14 Relaciones y características.

Por otro lado, la Figura 4:15, muestra la relación entre los conceptos definidos como <<Producto>> y el concepto denominado <<Valoración>>. Este último concepto es una valoración que los usuarios realizan sobre los productos. Para una mejor comprensión de las relaciones entre los conceptos se recomienda mirar la Sección 4.2 de la presente tesis doctoral. En consecuencia, la valoración del producto es, en sí misma, una característica del producto en relación con el usuario, ya que es el usuario quien se encarga de evaluar los productos ofertados. Por lo tanto, el concepto permite evaluar/calificar a un producto a través de un usuario.

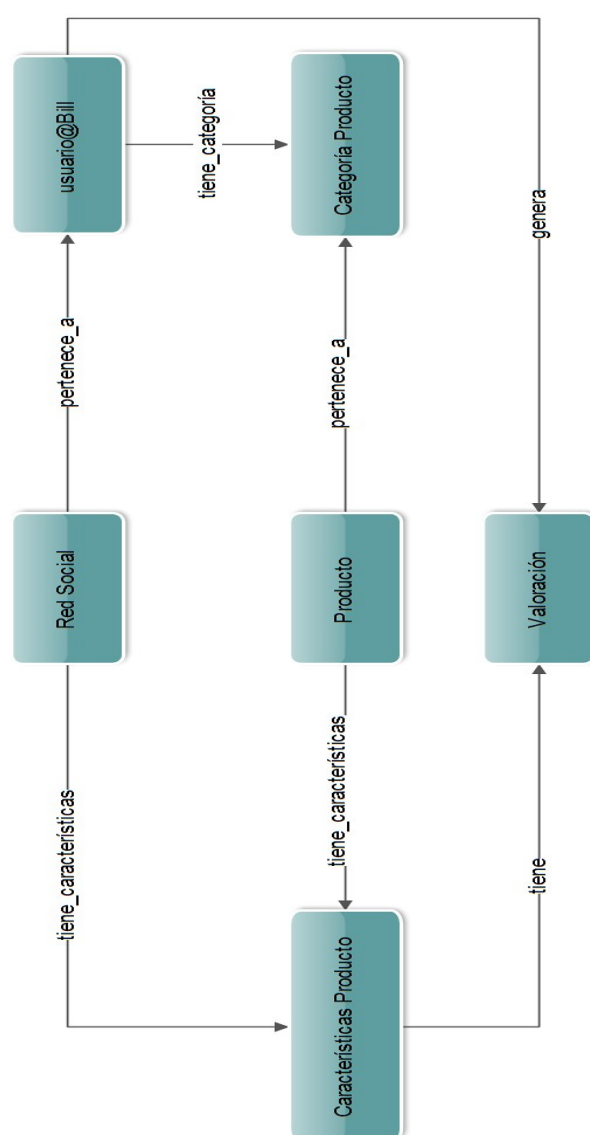


Figura 4:15 Relaciones de alto nivel para una valoración.

4.5.2. Semántica del modelo

No existe un significado bien establecido para el término de anotación semántica o información abstracta. Sin embargo, la anotación semántica es una nueva generación de esquemas de uso, el objetivo es permitir nuevos métodos de acceso a la información y ampliar los ya existentes [Kiryakov et al., 2011]. Además, existen herramientas de anotación que permiten añadir un contenido semántico (i.e., XML, OWL, RDF), estas herramientas permiten estructurar la información mediante una clasificación en base a los conceptos semánticos establecidos. De manera que para comprender la información almacenada en la web es necesario de un identificador único, capaz de distinguirlo, de forma que se le puedan asociar características, relaciones, aserciones lógicas o cualquier tipo de metainformación.

Por lo tanto, en la presente investigación se ha desarrollado una ontología para representar los criterios y conceptos relativos a productos ofrecidos en el mercado, sus características y las preferencias de los usuarios de las redes sociales, es decir, aquellos conceptos definidos en el modelo propuesto. Una anotación semántica se considera como una información sobre las entidades o conceptos que conforman la ontología. En consecuencia, como herramientas complementarias a las de anotación disponemos de los editores de ontologías, que facilitan la tarea de definición de las mismas, la unión de diferentes ontologías y el desarrollo distribuido de ontologías [Noy et al., 2001].

En el proceso es capaz de reunir el conocimiento desde las diversas fuentes de información que se encuentran disponibles en los medios sociales. En esta investigación se presentan dos casos de estudio descritos en la Sección 6.3 donde se utiliza la red social Twitter. Por lo tanto para este caso, información de la red social Twitter (a través del conector Streaming Twitter). El objetivo es anotar semánticamente una ontología a partir de la identificación de la información relevante publicada en las redes sociales por los usuarios.

En este sentido, la ontología sirve de soporte para un Sistema Basado en Conocimiento, que a su vez, se basa en la información capturada desde las redes sociales almacenada en el repositorio semántico. En consecuencia, a partir del modelo ontológico

se desarrollará una arquitectura para un sistema capaz de discernir entre los mejores productos y servicios que el usuario prefiera, la arquitectura se describe a detalle y presenta en la Sección 5.2 de la presente investigación.

Para realizar este procedimiento, se divide en cuatro pasos que permiten realizar el proceso de recabar la información sobre los distintos productos, marcas, sus características y servicios ofertados en las redes sociales. Además, el contenido sobre los productos, marcas, servicios, sus características y las preferencias de los usuarios se almacena en los conceptos establecidos en la base de conocimiento (Ontología).

Después del proceso para poblar la ontología se realiza la transformación difusa. En otras palabras, se lleva a cabo la teoría de conjuntos difusos, dicha teoría proporciona un marco para la representación de la incertidumbre de muchos aspectos del conocimiento humano. Por lo tanto, es necesario describir la teoría de conjuntos difusos, el motor semántico con lógica borrosa y sus aspectos importantes relacionados con la presente investigación.

4.6. Sumario

En este capítulo se ha diseñado un nuevo modelo conceptual multi-dominio para la representación del conocimiento desde distintos dominios a partir de los conceptos y criterios definidos en un modelo formal. Asimismo, se han definido los conceptos principales que permiten establecer relaciones y criterios para la caracterización del contenido y la representación del conocimiento.

Además, las comprobaciones que el modelo realiza sobre las distintas relaciones entre los conceptos definidos que representan la gestión del conocimiento, y las valoraciones que el Sistema Basado en Conocimiento procesa para establecer las características de un producto o servicio, tienen como resultado la capacidad semántica y de soporte del contenido estructurado en el diseño del modelo propuesto en esta investigación. Asimismo, es posible recoger la publicación, edición e intercambio de información que los usuarios realizan en los medios sociales.

El Capítulo 5 completa la semántica del modelo conceptual multi-dominio con el diseño de una arquitectura que permita validar el modelo conceptual. Por lo tanto, se describen los componentes que integran el diseño de un marco computacional de trabajo para la interpretación de la información. Asimismo, la descripción del proceso de recomendación a partir de una herramienta de recomendación que hace uso del contenido almacenado en la base de conocimientos para facilitar la generación de recomendaciones hacia los usuarios sobre productos, sus características y servicios ofertados en la Web.

Capítulo 5

Arquitectura de la solución

propuesta

Actualmente, en la Web 3.0 se posibilita la convergencia entre distintas formas de contenido, (web TV, portales web de contenidos, Protocolo Televisión por Internet, Televisión Digital Terrestre, portales de entretenimiento Web, Facebook, Twitter, etc.). Además, la información que se encuentra almacenada en las distintas plataformas de comunicación social, se considera como un área potencial para la representación del conocimiento mediante la utilización de las TS que han sido emergentes en la investigación en los últimos años.

En este capítulo se diseña y describe una arquitectura para un marco computacional de trabajo en la representación del conocimiento para múltiples dominios. El diseño del marco, es basado en el modelo multi-dominio propuesto en el Capítulo 4. Además, se describen cada uno de los distintos componentes que integran una arquitectura de soluciones al planteamiento inicial del problema en la investigación. Asimismo, en el diseño del marco computacional incluye una ontología que describe las relaciones de los conceptos definidos en el modelo conceptual descrito en el capítulo anterior, que es implementada como una base de datos. Una vez formalizado el marco se considera un motor de lógica borrosa para la generación de recomendaciones. En consecuencia, un sistema basado en conocimiento permite representar un contenido caracterizado a partir de la extracción de la información semántica de los medios sociales. Finalmente, el proceso de recomendación se realiza a partir de una herramienta de recomendación que utiliza el contenido caracterizado sobre los distintos productos, servicios y sus características que satisfacen las necesidades de un usuario aplicado a distintos dominios.

5.1. Descripción general

El modelo conceptual multi-dominio se basa en la representación de la información que se adquiere a partir de las fuentes de información disponibles en la Web, específicamente de las redes sociales. Dicho modelo conceptual multi-dominio se formaliza a través de una ontología que permite la gestión de la información, mediante la comprensión de un conjunto de entidades, relaciones y funciones sobre la conceptualización formal del modelo descrito en el Capítulo 4. Llegado a este punto de la investigación surge la necesidad de aplicar el modelo en un entorno real. Por lo tanto, es necesario plantear el diseño de un marco computacional de trabajo que, basado en el modelo conceptual, permita su aplicación y sirva para la representación del conocimiento, pudiendo además, sin cambiar sus conceptos principales del modelo, pueda ser trasladado a distintos dominios.

En consecuencia, se ha definido una arquitectura basada en el modelo propuesto, además se ha decidido representar la ontología que define el modelo conceptual multi-dominio en forma de una base de datos relacional. De esta forma, es definida en la arquitectura para la solución propuesta al proceso que permite el almacenamiento de los datos en las diversas instancias. Así, de esta forma se prevén problemas de rendimiento y escalabilidad según comentan los siguientes autores [Astrova, Korda, and Kalja 2007 ; Trinkunas and Vasilecas 2015 ; Jeong et al. 2006 ; Vysniauskas and Nemuraite 2015 ; Cuadrado 2009]. Por lo tanto, los datos extraídos son definidos y almacenados en las distintas instancias y clases de la ontología del sistema. Para esta conceptualización se emplea el lenguaje de ontología OWL, como un medio para expresar las características del modelo y, además, permite representar el conjunto de conceptos, propiedades y relaciones del modelo conceptual multi-dominio.

Dentro de este orden de ideas, se han definido distintos componentes que se integran en la arquitectura de una plataforma como marco de trabajo computacional para la representación del conocimiento. De esta forma, los componentes permiten establecer una arquitectura capaz de interpretar la información extraída a partir de los medios sociales

y, a su vez, permite la caracterización del contenido sobre los distintos usuarios, redes sociales y productos. En consecuencia, la información que proporciona el modelo sirve de base para un Sistema Basado en Conocimiento que permite representar la información para múltiples dominios. Además, la caracterización del contenido permite, a partir de los componentes establecidos, generar recomendaciones considerando el contenido caracterizado en el modelo conceptual multi-dominio y las preferencias obtenidas del comportamiento de los usuarios.

Por lo tanto, el modelo conceptual permite la caracterización de la información adaptable a distintos dominios de contenido estructurado y de conocimiento semántico. Asimismo, el diseño de la arquitectura propuesta está basado en el modelo descrito en esta investigación. Gracias al diseño planteado en la arquitectura, se aplica el modelo conceptual propuesto a distintos dominios. Por lo tanto, a continuación se describirá la propuesta del diseño de la arquitectura y cada uno de los componentes que integran el marco computacional de trabajo basado en el modelo descrito en el capítulo anterior.

5.2. Arquitectura para la representación del conocimiento

El diseño de la arquitectura es basada en el modelo multi-dominio propuesto en la presente investigación y descrito en el Capítulo 4. Además, la arquitectura integra distintos componentes que hacen del proceso una solución al planteamiento inicial del problema.

Para poder recomendar un producto a los distintos usuarios que demandan información personalizada a sus necesidades, es necesario realizar con anterioridad, un análisis sobre el contenido de los productos que permitan darle sentido e interpretar sus características. Por lo tanto, la creación de distintos sistemas basados en conocimiento para la generación de recomendaciones de productos, permiten explotar el área ofreciendo respuestas específicas a los usuarios. En esta investigación se hace uso de

una herramienta de recomendación para la generación de recomendaciones a partir de la información almacenada en el modelo propuesto.

Ahora bien, la definición de un modelo conceptual multi-dominio y las TS ayudan a generar y gestionar un contenido estructurado para distintos dominios en una arquitectura propuesta para un marco computacional. Asimismo, es necesario conocer los gustos de los usuarios que permitan establecer las preferencias de un usuario sobre aquellos productos relevantes que sirven de preferencia a partir de un filtrado de información semántica para la generación de recomendaciones personalizadas.

En la Figura 5:1, se muestra el esquema de la arquitectura para el diseño de un marco computacional basado en el modelo conceptual multi-dominio propuesto en esta investigación.

Los razonadores semánticos son aplicaciones computacionales que permiten realizar inferencias y generar conocimiento a partir de un conjunto de axiomas y hechos. Además, los razonadores utilizan un motor de inferencias y un conjunto de reglas que pueden estar expresadas en diferentes lenguajes.

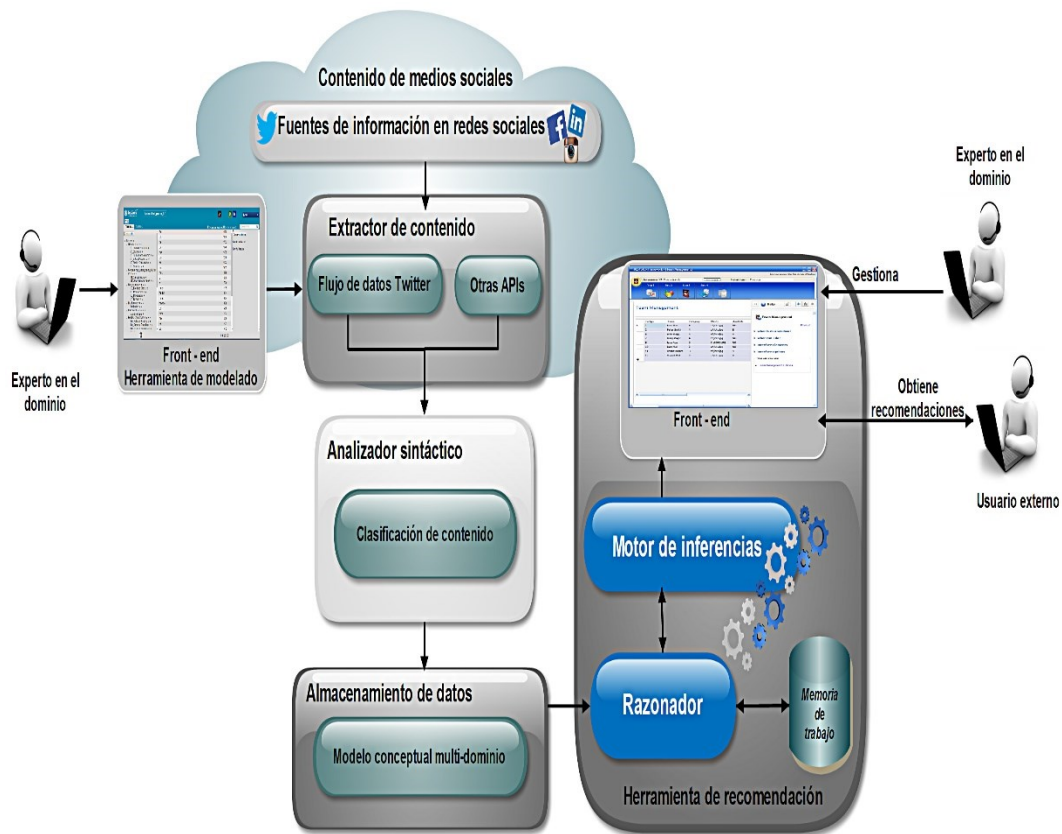


Figura 5:1 Arquitectura general del marco computacional basada en el modelo multi-dominio.

En el diseño de la arquitectura propuesta, antes de describir cada uno de los componentes que integran el marco computacional de trabajo, es necesario considerar los siguientes elementos de la arquitectura:

- Una base de conocimiento que contiene la definición del modelo propuesto permitiendo la caracterización de medios sociales, productos, marcas, servicios y perfiles de usuarios.
- Para la caracterización de un producto, un perfil y una red social es necesario partir de la información extraída y almacenada en la Web.
- En el proceso de recomendación puede intervenir un experto en el dominio. Los expertos pueden ser personas u otros sistemas que tienen

conocimiento detallado sobre un determinado dominio y vuelcan su experiencia y conocimiento en el marco computacional para poder realizar recomendaciones a partir de un SR. Asimismo, comprueban que los resultados del sistema son los correctos.

- Un motor de razonamiento basado en reglas permite clasificar la respuesta del sistema. En esta investigación se ha optado por un motor de reglas basado en lógica borrosa.
- Una memoria de trabajo permite almacenar la información de cada uno de los procesos de recomendación para cada usuario.
- Una interfaz de comunicación con los expertos y el usuario permitirá visualizar la recomendación dada por el sistema.

La arquitectura se basa en la representación de la información de múltiples dominios que se adquieren a partir de las distintas fuentes de información disponibles en la Web, específicamente de las redes sociales. Asimismo, la arquitectura está basada en el modelo ontológico descrito en el Capítulo 4. De tal forma que, para validar el modelo propuesto, es necesario el diseño de un marco computacional que integre dicho modelo y, además, distintos componentes que en su conjunto permitan hacer del proceso una solución al planteamiento inicial del problema propuesto en la presente investigación. Por lo tanto, a continuación se describe el diseño de un marco computacional basado en el modelo y en la información extraída de las fuentes disponibles y almacenadas en la Web. En esta investigación, se ha optado por utilizar la red social Twitter como una fuente para la extracción de la información.

Para llevar a cabo el proceso de extracción y clasificación de la información sobre los distintos sectores, en este caso sector hace referencia al concepto multi-dominio y, marcas o entidades hace referencia a los distintos productos, ambos son conceptos definidos en el modelo ontológico descrito en el Capítulo 4 y mostrados en la Figura 4:1.

Para este proceso de extracción de los tweets, se utiliza un crawler que extrae aquellos tweets (los 140 caracteres) que contienen las palabras definidas para los distintos sectores y marcas (hablar de sectores, es hacer referencia a los múltiples dominios, asimismo, el hablar de marcas es hacer referencia a los distintos productos). La extracción se realiza a través de un conector en Java que permite extraer el flujo de datos utilizando la API Streaming de Twitter.

Las palabras sobre los sectores y marcas del dominio a modelar se encuentran en una lista de términos. Dicha lista de términos es definida por un experto en el dominio, además existe una lista de twitteros que almacena los tweets que escriben cada uno de ellos. Por lo tanto, en la extracción se identifican aquellos tweets que contienen las palabras definidas en la lista de términos sobre el sector y marca a ser analizados. Una vez extraída la información se utiliza un analizador sintáctico (Parser), que clasifica el contenido a un formato de trabajo deseado. De manera que, una vez obtenidos los tweets con las palabras claves son clasificados, por un lado, los tweets que contienen las palabras deseadas y, por otro lado, características específicas, tales como, término, año, mes, día, usuario, etc.

Para generar la ontología propuesta y descrita en el Capítulo 4, se lleva a cabo a través de una extracción de tweets que permiten analizar los sectores definidos.

Una vez clasificada la información, se caracteriza y es almacenada en la base de conocimiento del modelo ontológico propuesto para esta investigación. De esta forma, el modelo ontológico almacena la información sobre las características de los sectores y marcas o entidades definidas (Dominios, productos, perfiles de usuarios, medios sociales). Además, la base de conocimiento está basada en la ontología propuesta en esta tesis doctoral. Así, la información caracterizada se transforma en un contenido relevante, de manera que dichos datos, permiten crear un nuevo conocimiento para ser modelado sobre los sectores, marcas y usuarios.

A continuación se describen cada uno de los elementos que componen una arquitectura para el diseño de un marco computacional basado en el modelo conceptual multi-dominio propuesto en esta tesis doctoral.

5.3. Componentes que integran el marco computacional de trabajo

Los componentes que integran el marco pretenden dar una solución al caracterizado de los medios sociales, productos y usuarios a través de los conceptos definidos en el modelo conceptual multi-dominio propuesto. En consecuencia, el marco computacional de trabajo debe poder caracterizar un contenido estructurado sobre los distintos productos, marcas, sus características y servicios ofertados en las redes sociales, a su vez, puede modelar y gestionar el conocimiento de diferentes perfiles de usuarios, productos y medios sociales caracterizados en múltiples dominios. Por lo tanto, la información almacenada debe servir para la generación de recomendaciones personalizadas.

5.3.1. Herramienta de modelado

Este componente se define como una herramienta de modelado que permite modelar los productos, marcas, sus características y servicios ofertado en las redes sociales, a su vez, analizar los datos extraído de Twitter. Asimismo, la herramienta gestiona el conocimiento a través de conceptos de fácil comprensión para el usuario y, además se basa en el modelo conceptual propuesto en el Capítulo 4.

La manera de intercambiar la información sobre un determinado dominio que se encuentra almacenado en la base de conocimiento, se realiza a través de la herramienta de modelado. Por un lado, la herramienta de modelado facilita la gestión de la información. Por otro lado, un experto en el dominio proporciona su conocimiento en el dominio a través

de reglas que modela por medio de la herramienta. La Figura 5:2, muestra la herramienta de modelado para la gestión de la información.

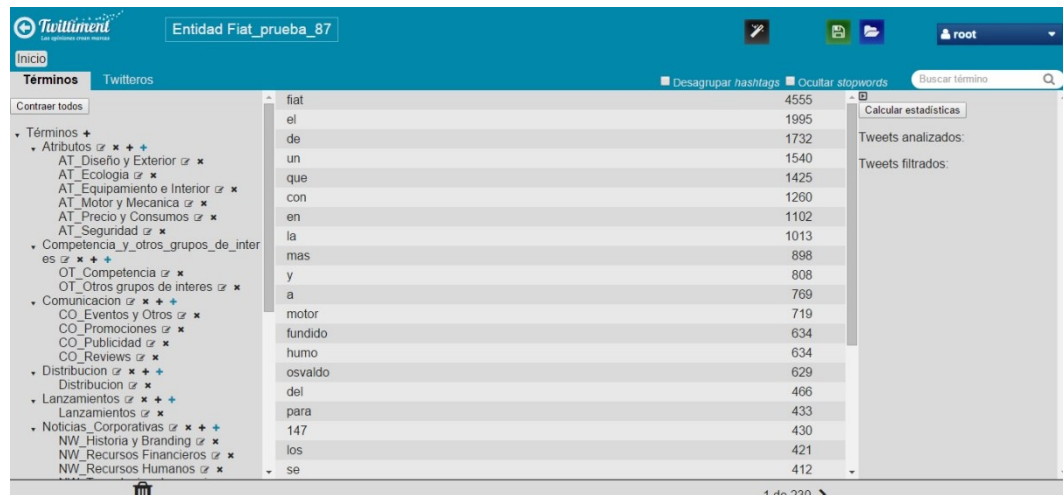


Figura 5:2 Herramienta de modelado para la gestión de la información

En este sentido, la función del experto es definir las características relevantes sobre productos del dominio que sean analizados a través de los conceptos y criterios definidos en la base de conocimiento (ontología del modelo). Por lo tanto, el experto especifica semánticamente las características relevantes de los productos utilizando la herramienta de modelado, además crea las correspondientes reglas borrosas hacia los productos del dominio que ayudan a la toma de decisiones del sistema. La Figura 5:3, muestra como el experto utiliza la herramienta de modelado para la definición de los conceptos para la extracción de la información.

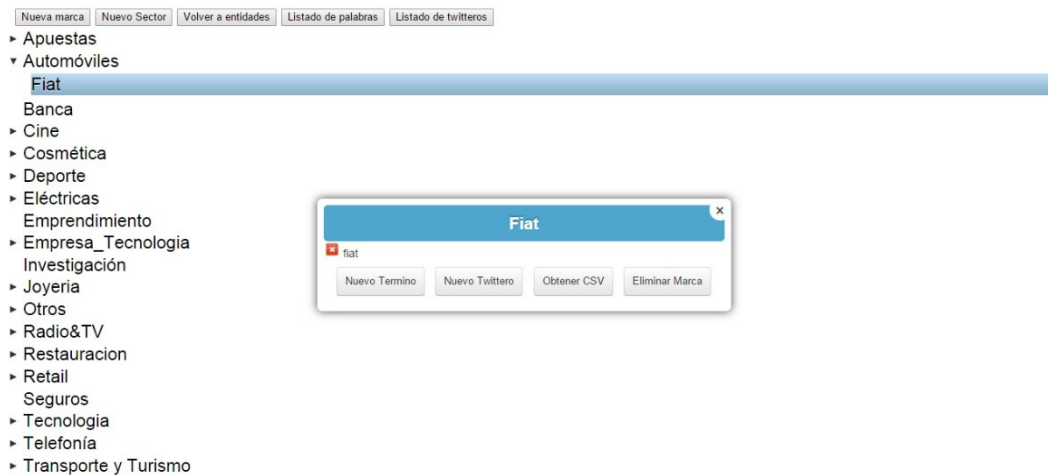


Figura 5:3 Definición de criterios y conceptos para la extracción mediante la herramienta de modelado

En la base de conocimiento, por un lado se encuentran los datos que un experto en el dominio del problema utiliza para la toma de decisiones. Por otro lado, el contenido extraído de los medios sociales que mantiene un tipo de relación con los datos del dominio. Además, es necesario un motor de inferencia para elegir la decisión a tomar partiendo de las reglas y los conocimientos adquiridos. Finalmente, como complemento integrado en la arquitectura, una herramienta de recomendaciones que de explicación de las decisiones que ha tomado el sistema. Por lo tanto, en las siguientes secciones se describen los pasos y módulos a seguir que integran el proceso de recomendación, así como la conexión que existe entre cada uno de ellos. Por consiguiente, dichos módulos permiten alcanzar la generación de las recomendaciones a partir de la arquitectura propuesta. A continuación se describen los pasos necesarios al utilizar la herramienta de modelado.

- El primer paso es identificar la fuente de información a partir de los medios sociales en la Web. En consecuencia, se obtiene información de las redes sociales, aplicaciones que pertenecen a los medios sociales en Internet. Son estructuras compuestas por un conjunto de actores, tales como individuos u

organizaciones, lo cuales están relacionados de acuerdo a distintos criterios y conceptos (relación profesional, amistad, parentesco, etc.). Por lo tanto, se extrae una información que sirve para identificar los productos, marcas y servicios que son mencionados a través de la red, así como las características y preferencias que los usuarios mantiene hacia los productos, marcas y servicios ofrecidos en el mercado, de manera que se establece un perfil a partir de la información caracterizada y anotada semánticamente en la base de conocimiento con respecto a los conceptos y criterios establecidos.

- Un segundo paso es el recolector de información, ya que una vez identificada la información es necesario extraer la información mediante un algoritmo de búsqueda que permite identificar los temas requeridos por el sistema. Para esta investigación se trabajó con la red social Twitter. Sin embargo, el proceso permite extraer información desde distintos recursos ofertados en la Web.
- El tercer paso es clasificar la información relevante que es extraída y transformada al formato adecuado. Asimismo, realizar las operaciones correspondientes para poblar la base de conocimientos a partir de los conceptos y criterios establecidos en la ontología del modelo.
- Finalmente, el cuarto paso es la base de conocimiento que almacena la información detallada sobre las características de los usuarios, redes sociales y productos, además de los criterios y conceptos establecidos por expertos que se relacionan con la información sobre un dominio mediante una herramienta de modelado. Por lo tanto, la base de conocimiento utiliza una ontología que gestiona los criterios y conceptos del contenido relevante, además permite especificar una estructura hacia un sistema a partir del modelo conceptual propuesto en esta investigación.

5.3.2. Fuentes de información en redes sociales y extractor de contenido

Es un componente que contiene a las distintas aplicaciones que pertenecen a los medios sociales en Internet, las cuales permiten conectar entre usuarios y extraer información. Es decir, estructuras compuestas por un conjunto de usuarios, tales como individuos u organizaciones, lo cuales están relacionados de acuerdo a distintos criterios y conceptos (relación profesional, amistad, parentesco, etc.). En este caso, el sistema utiliza la red social Twitter como una fuente para la extracción de la información que sirve para identificar los productos, marcas y servicios que son mencionados por los usuarios a través de la red, así como las características y preferencias que los mismo usuarios mantiene hacia aquellos productos, marcas y servicios ofrecidos en el mercado.

La información se encuentra accesible en el contenido de medios sociales disponible en la Web. En este sentido, el contenido de medios sociales se refiere a las plataformas de comunicación en línea donde el contenido de productos, marcas, servicios, sus características y preferencias de los actores es creado por los propios usuarios mediante el uso de las tecnologías de la Web que facilitan la edición, la publicación y el intercambio de información, ya sea por conversaciones de textos, fotografías, imágenes, videos o audio. Además, como ya se ha comentado en el Capítulo 2, los medios sociales pueden adoptar muchas formas diferentes, incluyendo foros de Internet, Blogs de personalidades, blogs sociales, wikis, podcasts, fotos y vídeo.

Por otra parte, decir que la extracción de los datos para generar un contenido valioso en las redes sociales es cada vez más usado por las empresas y usuarios, ya que permiten alcanzar metas y objetivos planteados en tiempos ventajosos. Por lo tanto, es necesario extraer la información y representarla de acuerdo al modelo conceptual multi-dominio propuesto en la presente investigación. Para ello, existe una normalización del conocimiento, es decir un proceso de estructuras para la organización del contenido, que puede ser a través de la creación de ontologías que hacen posible llevar a cabo el proceso mediante una semántica de contenidos. La Figura 5:4 muestra el proceso genérico para un análisis en la extracción del contenido desde distintos paradigmas.

Para la extracción existe un componente denominado analizador sintáctico, también conocido como Crawler, su función principal del componente es la sustracción de la información semántica de las diversas fuentes de información en los medios sociales y, de los distintos productos así como sus características que son relevantes para los usuarios que demandan la información.



Figura 5:4 Proceso genérico para un análisis en la extracción de contenidos.

Con la extracción de los datos es posible medir de forma cuantitativa lo que está ocurriendo en las redes sociales, de manera que un componente de análisis de datos proporciona un significado personalizado y relevante para el usuario a través del sistema. Por lo tanto, el componente analizador de contenido es un recolector de información que mediante un algoritmo de búsqueda, para este caso, mediante un conector en Java permite extraer el flujo de datos utilizando la API Streaming de Twitter. En este sentido, el conector identifica los temas requeridos por el sistema. Sin embargo, la arquitectura propuesta para el diseño del marco se puede ampliar para la búsqueda de distintos recursos ofertados en la Web.

En consecuencia, el componente extrae el contenido sobre productos, marcas, servicios, sus características y las preferencias de los usuarios a partir de las menciones

que realizan mediante la red social. En este sentido, un experto en el dominio define la información que se va a extraer sobre los productos, marcas, sus características y servicios. Por lo tanto, el componente permite detectar las fuentes de información almacenadas en la Web. Una vez detectada la información, realiza una selección de información relevante que va en función a los distintos criterios de interés establecidos por el usuario y el experto.

5.3.3. Analizador sintáctico

Para el diseño del marco es necesario realizar una clasificación del contenido relativo a productos, marcas, servicios ofrecidos en el mercado, sus características y las preferencias de los usuarios de las redes sociales. Por lo tanto, se hace uso de un analizador sintáctico para clasificar la información relevante y transformarla al formato adecuado. En esta investigación, se utiliza un formato CSV (del inglés comma-separated values), dicho formato permite representar datos en forma de tablas separar los diferentes campos por comas, lo que es útil en la clasificación de la información. Se ha elegido este formato para la clasificación de información debido a su simplicidad y facilidad de conexión con el resto de componentes de la arquitectura diseñada. Un ejemplo del formato se puede ver en la Figura 5:5.

	A	B	C	D	E	F	G	H	I	J	K	L
1	Marca;	Término de búsqueda;	ID Tweet;	Fecha;	ID Twittero;	Nombre;	Apodo;	Followings;	Followers;	Tweet;	Es RT;	ID Twittero
2	audi;	audi;	481557452987457536;	Wed Jun 25 00:00:18 CEST 2014;	609343327;	Graciela Alonso.	SatanicEight;	564;	594;	En est		
3	audi;	audi;	481557473187622913;	Wed Jun 25 00:00:23 CEST 2014;	167634365;	Jose david;	sejo1986;	191;	67;	RT @machecor:		
4	audi;	audi'	familia anormal de mierda;	-1;	-1;	-1						
5	audi;	audi'	Audi A8 y sus 5.1 metros de lujo.	http://t.co/iyFSM3QaFw;	0;	1690137476;	373861737457393664					
6	audi;	audi;	481557570088624128;	Wed Jun 25 00:00:46 CEST 2014;	181276161;	JD Gomez;	Jdon360;	465;	203;	RT @machecor:		
7	audi;	audi;	481557819616149505;	Wed Jun 25 00:01:45 CEST 2014;	136395728;	Ivan Puentes;	pillin182;	292;	16;	CLAUDIABAH		
8	audi;	audi'	en un grup en el ultimo minuto se clasifica Grecia;	0;	224280985;	481555865263431680						
9	audi;	audi'	de la Venezuela honda. La estudia a fondo;	0;	1653801763;	481525339147866112						
10	audi;	audi'	si salimos campeones bonificacion del 1%;	-1;	-1;	-1						
11	audi;	audi;	481558110444589058;	Wed Jun 25 00:02:55 CEST 2014;	498822915;	1 año juntos 020213	gattuso_96;	424;	312;	Fol		
12	audi;	audi;	481558230255271936;	Wed Jun 25 00:03:23 CEST 2014;	146935563;	Amarilla	LizHenao;	179;	578;	RT @machecor:		

Figura 5:5 Ejemplo del formato CSV para la clasificación del contenido.

En la clasificación del contenido, se obtienen los criterios tales como, productos, marcas, servicios, sus características, el Id-tweet, el texto del tweet, el número de seguidores, el nombre del tweet, entre otros. La información es usada para ser almacenada en la base de conocimiento del modelo ontológico.

5.3.4. Almacenamiento de datos

La base de conocimiento almacena información detallada sobre las características de los usuarios y productos, además de los conceptos relacionados con el modelo conceptual propuesto y descrito en la Sección 4.3. A partir de la definición de una ontología se ha creado una base de datos para el almacenamiento de los datos sobre productos, sus características puestos en esta investigación.

Por lo tanto, la base de conocimiento permite especificar la estructura del modelo propuesto en esta tesis. La descripción de alto nivel del modelo de datos de la ontología se compone de cuatro pilares principales: el modelado de un producto; las preferencias que están establecidas por un usuario; modelado de perfiles de usuario y el pilar de medios de comunicación social de donde se extrae la información. La descripción de alto nivel del modelo de datos se presenta en la Figura 5:6.

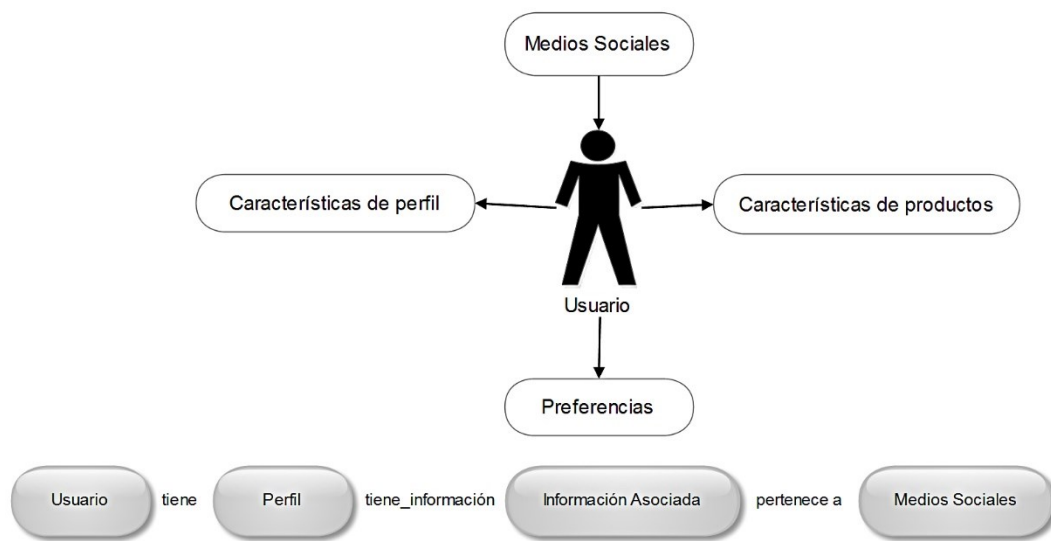


Figura 5:6 Modelado de alto nivel para gestión de la información.

Para la representación del conocimiento, los pilares de alto nivel del modelo de datos de la ontología propuesta se describen a continuación:

- **Usuario.**- Se refiere a los actores que son identificados a través de un Id.
- **Medios Sociales.**- Las redes sociales para su caracterización.
- **Productos caracterizados.**- Marcas que se almacenan en la Web y que mantienen sus propias características.
- **Perfiles caracterizados.**- Diferentes perfiles caracterizados y vinculados con productos y redes sociales.
- **Preferencias.**- La clase establece las preferencias del usuario hacia los productos ofertados en la Web a través de reglas establecidas por el sistema.

En la construcción de la ontología permite gestionar el conocimiento sobre la base de los conceptos del modelo conceptual multi-dominio propuesto en esta investigación. Un modelado de gustos, perfiles, productos y redes sociales que se genera entre ellos

permite generar una caracterización de la información. Así, la información se transforma en un contenido relevante, de manera que dichos datos, permiten crear un nuevo conocimiento sobre los usuarios.

Por otro lado, el componente gestiona y almacena de forma estructurada la información representada explícitamente por el modelo conceptual propuesto, en consecuencia, el componente filtra la información para obtener una búsqueda simple de la información relevante, que a su vez, permita representar un contenido estructurado claro y preciso para la caracterización de los siguientes recursos: redes sociales, productos y perfiles de usuario. Un ejemplo del proceso de la gestión del conocimiento en el sistema se puede apreciar en la Figura 5:7.



Figura 5:7 Proceso genérico para la gestión del conocimiento.

A medida que un usuario valora las preferencias previas sobre ciertos productos ofertados en la Web, el componente tiene el fin de transferir el conocimiento, valorarlo y representarlo.

Una forma de realizar la implementación de la base de conocimiento que recoja la definición de los conceptos del modelo formal, es mediante la realización de operaciones

correspondientes sobre ellas. Por lo tanto a continuación se describen las herramientas empleadas para la representación del conocimiento.

5.4. Proceso de recomendación

Para el proceso de recomendación, se hace uso de una herramienta de recomendaciones que, sobre la base de los componentes antes descritos para la interpretación del contenido y, mediante el filtrado de información semántica, debe proporcionar una respuesta a las necesidades de un usuario. Por consiguiente, como se ha detallado anteriormente, se asume que la información se extrae de los medios de comunicación social. Así, a partir de la definición de una ontología se ha creado una base de datos relacional que almacena el contenido de productos, sus características y servicios ofertados en la Web. Al mismo tiempo, el sistema de recomendación incorporado en la arquitectura de la plataforma, permite gestionar el conocimiento de los expertos en un dominio, así como las solicitudes de un usuario externo. De esta manera se pueden tomar decisiones de acuerdo con la habilidad y la interacción con el usuario.

La recomendación se realiza a partir del cruce de las características recolectadas de los perfiles de usuario y las características de los productos de un dominio. Por un lado, los perfiles se obtienen a partir las diferentes menciones que el usuario realiza través de la red social Twitter para un dominio. Por otro lado, las características de los productos se definen a partir de los conceptos establecidos en la ontología que se alimenta a través de los Tweets publicados en un dominio.

Sobre la base del conocimiento extraído de los productos y los perfiles de usuario, el marco de trabajo es capaz de proporcionar recomendaciones hacia un dominio concreto, las recomendaciones son basadas en el contenido extraído de la red social Twitter que mejor se ajustan a un perfil de usuario.

En consecuencia, el marco computación del trabajo permite caracterizar el perfil de cada usuario y caracteriza los productos a partir de características definidas en los

conceptos de la ontología del sistema. Por otro lado, la ontología es poblada a través del API de Twitter para extraer los tweets de la red social que permite representar el conocimiento en un dominio. La funcionalidad del sistema propuesto se presenta en la Figura 5:1 de la Sección 5.2.

A continuación, se describen cada una de las fases que permiten llevar a cabo la recomendación personalizada mediante un marco computacional de trabajo propuesto en los medios sociales. Además, el marco de trabajo es basado en el modelo conceptual propuesto en la presente investigación.

- En la primera fase, la ontología permite definir la información que es necesaria recoger de los usuarios y los productos. La ontología establece conceptos para la caracterización de la información a partir de las menciones que el usuario realiza en los medios sociales. Para este caso, utilizamos la red social Twitter que contiene la información sobre diferentes dominios y almacenarla en la base de conocimiento del sistema. La extracción se realiza por medio de una herramienta de modelado que permite analizar los datos en Twitter. Además, la herramienta se basa en el modelo conceptual propuesto en el Capítulo 4 que gestiona el conocimiento a través de conceptos de fácil comprensión para el usuario. El modelo propuesto para esta tesis se detalla en la Sección 4.4.
- En la segunda fase, en el marco computación del trabajo se establece el perfil de usuario a partir de la información caracterizada y anotada semánticamente en la base de conocimiento. Además, alimenta la base de productos a partir de la información recogida de la red social para un dominio.
 - Por un lado, el sistema recoge la información de los intereses que menciona un usuario hacia determinados productos de un dominio en Twitter mediante el componente de extracción del sistema y la API de la red social Twitter.

- Por otro lado, en esta segunda fase. El marco a través del componente de extracción y la API de Twitter alimenta la ontología de productos a través de la información en los Tweets que considera relevantes para un dominio.
- Una vez que se ha establecido el perfil del usuario, se genera una tercera fase en la que a través del componente difuso se establece que productos para el dominio analizado cumplen una serie de reglas borrosas. Es decir, sobre la base de las reglas difusas definidas en el sistema se genera una evaluación sobre las características de los productos de un dominio. Estos productos se han descrito anteriormente y anotado semánticamente de acuerdo con criterios establecidos por expertos en la ontología del sistema y cargados en la base de productos mediante el componente de extracción (Fase 1 y Fase 2).
- Finalmente, una cuarta fase es la recomendación personalizada de productos que mejor se adapten a las características de un perfil de usuario sobre un dominio. La recomendación se realiza a partir del cruce de las características recolectadas del perfil de usuario y las características de los productos mencionados. Para dar prioridad a las recomendaciones sobre las características de un perfil, en esta investigación se caracterizan las menciones como una preferencia positiva y objetiva a través del usuario. Por ejemplo, si el usuario habla de <<seguridad>>, las reglas del sistema caracterizan que un usuario le gusta la seguridad. Por otro lado, el marco computación del trabajo también selecciona los productos a partir de las preferencias mencionadas por los usuarios. Por lo tanto, la selección de la información se realiza en base a las preferencias que el usuario menciona en el contenido extraído desde Twitter.

Antes de que el sistema pueda ser usado, el experto del dominio debe especificar semánticamente las características relevantes de los productos mediante los conceptos definidos en la ontología del sistema, además debe crear las reglas borrosas correspondientes con los productos del dominio.

Por lo tanto, la inferencia del razonamiento difuso se almacena en un componente llamado memoria de trabajo, el componente es una base de datos. La información se utiliza para generar futuras recomendaciones a usuarios externos sobre las diferentes preferencias almacenadas en la base de conocimiento. La Figura 5:8, se muestran los pasos a seguir durante el proceso de la recomendación que se realiza en el sistema.



Figura 5:8 Pasos en el proceso de recomendación.

En este sentido, el marco computacional de trabajo, por un lado se encarga de representar un conocimiento sobre productos, sus características y servicios en un dominio a partir de los conceptos definidos en el modelo propuesto. Por otro lado, permite conectar con el usuario externo que accede al sistema mediante una cuenta asociada a Twitter que le permite el acceso al contenido del sistema y, que a su vez, permite conocer sus preferencias.

Para facilitar las recomendaciones en el sistema, por un lado es necesario conocer el perfil del usuario a través de sus preferencias, y por otro las características de los productos de forma que se permita establecer una relación con el perfil de usuario. Para facilitar esta relación, el sistema hace uso de un componente difuso que permite establecer reglas que transforman valores continuos en variables discretos usando modificadores lingüísticos, tal y como se ha descrito en el estado de la cuestión en la Sección 2.4.4 del

Capítulo 2. De esta manera, el componente difuso permite evaluar las características de un producto a partir de las menciones del usuario dentro de la red social.

Una vez conocida la información de la caracterización de los productos de un dominio, se realiza una clasificación de los productos asociados con las normas que permiten una caracterización de los productos reales y la identificación de los tipos de categorías en los diferentes dominios. La ontología del sistema permite gestionar la información obtenida de una clasificación jerárquica de las categorías. Con esta finalidad, la obtención de una información detallada que permite generar la recomendación hacia el usuario. En relación con las implicaciones el contenido se infiere de la información almacenada en las redes sociales.

Después de describir los pasos a seguir y los módulos que integran el proceso de recomendación, a continuación se describe la herramienta de recomendación incluida en esta investigación.

5.4.1. Herramienta de recomendaciones

Para la validación de la investigación realizada en esta tesis doctoral, se ha incluido un componente de recomendaciones en el diseño de la arquitectura, el componente es una herramienta de recomendaciones de ámbito genérico. Es decir, que permite representar cualquier conjunto de productos que posean características descriptibles. Por lo tanto, esta herramienta permite la generación de recomendaciones de contenido estructurado sobre productos, marcas, servicios y sus características a partir de las preferencias que mantienen los usuarios a través de las redes sociales. El motor de inferencias en el que se basa la herramienta trabaja con lógica difusa. Es decir, le llegan unas características entrantes que traduce en otras más técnicas y pondera que productos son los más acordes a la petición. Este proceso se describe más adelante, sin embargo a continuación se describe con mayor detalle el razonador y el motor semántico que utiliza la herramienta de recomendación.

5.4.1.1. Razonador basado en reglas fuzzy logic

Un razonador puede llevar a cabo procedimientos automáticos para inferir y generar nuevas relaciones. Existe un amplio rango de razonadores automatizados disponibles para la inferencia del conocimiento [Horridge and Patel-Schneider 2009]. Por lo tanto, los razonadores son aplicaciones computacionales que permiten realizar inferencias y generar conocimiento a partir de un conjunto de axiomas y hechos.

Para este componente, se ha optado por definir un razonador basado en reglas definidas con lógica borrosa, ya que este tipo de razonadores permiten la combinación del conocimiento de un experto y el conocimiento extraído a partir de datos experimentales para generar soluciones desde diversos problemas, según [Serrano-Guerrero et al. 2011; Alonso, Munoz, and Botia 2008]. En este sentido, el razonamiento difuso se construye mediante el uso de reglas para la generación de un controlador difuso. Es decir, este componente normaliza los valores de entrada al sistema. Usualmente se reducen a un intervalo de (A y B), donde A y B son variables propuestas para el modelado en un sistema. Asimismo, los valores de entrada después de ser normalizados son fuzificados. Por lo tanto, cada valor se pasa por las reglas definidas en la base de conocimiento y sobre las variables difusas, de manera que, las variables reales se transforman en valores difusos. Además, el componente controla la forma en que se ejecutan las reglas difusas del sistema.

El marco computacional de trabajo emplea un sistema de recomendación basado en reglas difusas de tipo Mamdani descritas en la Sección 2.4.4 del Capítulo 2.

En primer lugar, una serie de reglas difusas han sido descritas que permiten identificar los productos. Un ejemplo de este tipo de reglas empleado en un caso de estudio para la validación mostrado en la Sección 6.3.

Para definir los vehículos de acuerdo a sus características, se utilizaría las siguientes reglas:

- (1) Si *consumo es bajo* Y *precio es económico* Y *diseño es compacto* Entonces el *coche es urbano*.
- (2) Si *consumo es alto* Y *precio es caro* Y *combustible es gasolina* Entonces el *coche es deportivo*.

Para ser capaz de obtener el conjunto completo de reglas, se consideraron todas las combinaciones de reglas teóricamente posibles P_t , teniendo en cuenta el número de antecedentes p y el número de conjuntos difusos de entrada A_p considerado para cada antecedente. Así, para cada consiguiente, el número teórico de posibles reglas es:

$$(3) \quad P_t = \prod_n A_n \text{ for } n = 1 \dots p;$$

Por lo tanto, el componente controla la forma en que se ejecutan las reglas difusas del sistema. Además, el componente se comunica con una memoria de trabajo que almacena la información de cada uno de los procesos de recomendación y con el almacenamiento de datos sobre el contenido de productos, marcas, servicios, sus características y las preferencias de los usuarios de la red social para cada usuario.

5.4.1.2. Motor de inferencias

La teoría de conjuntos difusos y la lógica borrosa propuesta por [Zadeh, 1965], se reconoce como un formalismo adecuado que permite capturar el conocimiento impreciso y vago. La importancia de la lógica difusa se deriva del hecho de que la mayoría de los modos de razonamiento humano y especialmente razonamiento de sentido común, son aproximados en la naturaleza a través de funciones definidas en conjuntos difusos [Bobillo et al., 2009]. De forma que varias funciones de pertenencia se pueden utilizar en la

definición de un conjunto difuso. Algunas de las más utilizadas son la triangular y la función trapezoidal [J & Bo, 2008]. Además, una de las características más importantes de la lógica difusa es su capacidad para llevar a cabo el razonamiento aproximado [Schockaert, Victor, & Houben, 2006].

Los Sistemas basados en reglas difusas tienen algunas ventajas sobre otros formalismos: proporcionan una representación natural para el conocimiento humano, así como un modelo muy interpretable (desde la semántica de las reglas pueden ser fácilmente entendidas incluso para usuarios no expertos), son más sencillas, económicas y más robustos que sus versiones claras y, por último pero no menos importante, han demostrado comportarse muy bien en aplicaciones prácticas [Bobillo et al., 2009]. Por otra parte, un motor semántico de lógica difusa es un razonador que trabaja en la información almacenada en el repositorio semántico que sea capaz de discernir entre los mejores productos y servicios que el usuario prefiera.

Por lo tanto, en la presente investigación para el proceso de recomendación se basa en un sistema de lógica difusa. Además, las recomendaciones personalizadas se basan en el conocimiento extraído de la red social y se almacena en el modelo semántico. Por lo tanto, el componente de lógica difusa permite definir reglas para enlazar, por un lado la información almacenada en la base de datos y, por otro lado con la memoria de trabajo asociada entre las características de los usuarios y las características de los productos. De este modo, el conocimiento inferido por el sistema permite la generación de la recomendación personalizada para los usuarios.

5.4.1.3. Interfaz de usuario (Front-end)

Una interfaz de usuario permite la visualización de las respuestas del sistema. El componente integrado en la arquitectura sirve para conocer, por un lado, las relaciones entre características y su peso asociado con el producto que serán descritos por un experto en el dominio a representar, que a su vez, irá afinando las recomendaciones y haciendo más fiable el sistema. Por otro lado, generar las recomendaciones sobre los

contenidos de los productos y sus características, la herramienta los recoge de la información almacenada en la Base de conocimiento o almacenamiento de datos. La Figura 5:9 muestra la interfaz de recomendación, la pantalla principal que utiliza la herramienta para mostrar las diferentes acciones que los usuarios podrán realizar en su estatus actual.

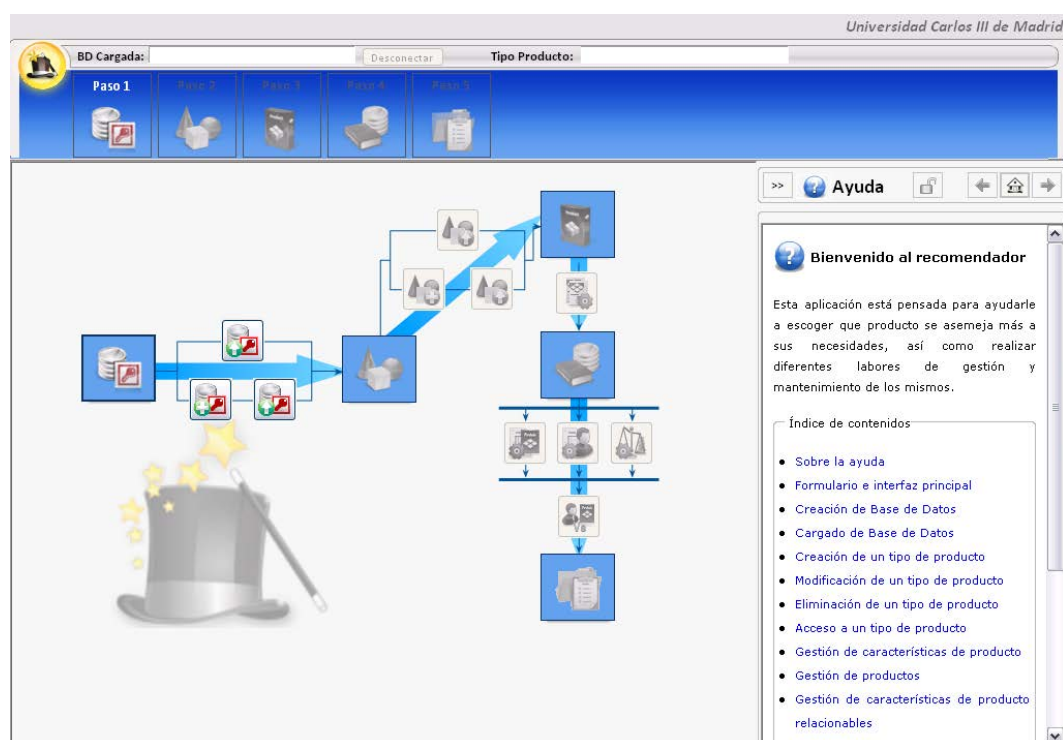


Figura 5:9 Herramienta de recomendación

Este componente realiza el proceso de los expertos y las solicitudes de un usuario externo. La recomendación se realiza a partir del cruce de las características recolectadas de los perfiles de usuario y las características de los productos de un dominio. Este proceso de las recomendaciones se detalla en la Sección 5.4.

Además, en el diseño propuesto para la arquitectura se establece dentro de esta, una interfaz de comunicación con un usuario externo. En la Figura 5:10 se muestra la interfaz que permite visualizar la respuesta del sistema, generando así, una recomendación personalizada hacia el usuario.

Solicitud y Resultado de Recomendaciones para Automoviles

AMBITO: Automoviles RECOMENDACIONES: 5

Datos Petición

CARACTERISTICA	VALOR
T_DISEÑO	(T)
T_PRECIO	(Alta)
T_SEGURIDAD	()

Productos Recomendados

NOMBRE	CATEGORIANCAP	FOTO	MODOE	LIBRES	DEPORTIVO	GPS	SITUACION	ESP	COD	DESCRIPCION
Opel Insignia 140CV Excellence	3	./Fotos/Opel/Insignia.jpg	SI	10	SI	SI	TENERIFE	SI	1	Nuevo Opel Insignia Edicion 2014.
Honda Civic 120CV Sport	3	./Fotos/Honda/Civic.jpg	SI	5	SI	SI	TENERIFE	SI	2	Honda Civic 2015 acabado Sport

Figura 5:10 Interfaz de recomendación de productos

La interfaz permite visualizar la respuesta del sistema hacia el usuario mediante una recomendación dada sobre un producto, marca o servicio y, a su vez, permite la interacción con el usuario externo para conocer sus preferencias y necesidades.

El usuario externo accede al sistema mediante una cuenta asociada a la red social Twitter que le permite el acceso al contenido y, que a su vez, permite conocer sus preferencias, de manera que se establece un perfil a partir de la información caracterizada y anotada semánticamente en la base de conocimiento con respecto a los conceptos y

criterios establecidos. Además, alimenta la base de productos a partir de la información recogida de la red social vinculada al usuario para un dominio.

5.5. Sumario

En este capítulo, se ha descrito el diseño de una arquitectura de soluciones en medios sociales mediante la representación del conocimiento para un marco computacional de trabajo. La arquitectura propuesta para la representación del conocimiento se basa en el modelo ontológico descrito en el Capítulo 4. Asimismo, se han definido y descrito cada uno de los componentes que integran una arquitectura en el diseño de un marco computacional. De manera que, el marco basado en el modelo propuesto de esta investigación brinda un soporte en la solución sobre los contenidos en la Web, medios sociales y usuarios, mediante el filtrado de información semántica y de contenido estructurado.

Por otra parte, en el proceso hacia una recomendación se ha descrito la herramienta que genera las recomendaciones sobre la base de los componentes descritos en la arquitectura, la herramienta utiliza la información que se encuentra almacenamiento de datos basado en el modelo conceptual propuesto y, a su vez, permite establecer a detalle cada uno de los recursos y conceptos que se emplean en el proceso con el fin de alcanzar una respuesta clara y precisa para las necesidades de los usuarios. Por lo tanto, el proceso hacia la generación de recomendaciones basado en el modelo conceptual propuesto en esta investigación provee una solución al planteamiento inicial del problema.

En este sentido, el diseño de la arquitectura de cada elemento mantiene una estrategia de selección sobre las decisiones que deben ejecutarse y que el razonamiento difuso debe implementar. Por lo tanto, la arquitectura propuesta permite evaluar el modelo multi-dominio en distintos dominios a partir de la información extraída de los medios sociales en un entorno computacional real, siendo Twitter la red social usada para la extracción de la información en esta investigación.

El Capítulo 6, describe la validación del modelo propuesto en esta tesis, en esa sección se integran las herramientas propuestas que permiten facilitar la adquisición del conocimiento y alcanzar los objetivos planteados al inicio de la presente investigación.

Asimismo, se presentan dos casos de estudio que permiten validar el modelo conceptual multi-dominio propuesto y descrito en el Capítulo 4. Finalmente, la comprobación de cada una de las hipótesis planteadas al inicio de la investigación se resalta.

Capítulo 6

Validación

Este capítulo describe el proceso de validación que se ha llevado a cabo a partir de las hipótesis planteadas. Para la validación se ha empleado la herramienta de modelado y un sistema basado en conocimiento. A partir de estas implementaciones, dos casos de estudio han sido desarrollados que permiten conocer, por un lado, la validación del modelo propuesto en esta tesis específicamente, para la representación del conocimiento sobre los distintos productos y servicios ofertados en los medios sociales. Por otro lado, en la validación se han demostrado las hipótesis planteadas al inicio de la investigación en un entorno de computación real.

6.1. Introducción

Una vez que se ha representado el conocimiento que permite resolver el problema a partir del modelo propuesto, en este capítulo se presentan dos casos de estudio que permiten validar, por un lado, la representación del conocimiento del modelo conceptual.

Por otro lado, la comprobación de las hipótesis planteadas al inicio de esta tesis doctoral.

En este sentido, el modelo conceptual propuesto y definido en el Capítulo 4 de esta tesis. Debe poder caracterizar un contenido estructurado sobre los distintos productos, marcas, sus características y servicios ofertados en las redes sociales, a su vez, puede modelar y gestionar el conocimiento de diferentes perfiles de usuarios, productos y medios sociales caracterizados en múltiples dominios. Por lo tanto y en consecuencia, a partir de la representación del conocimiento del modelo proyectado en dos casos de estudio se comprueban cada una de las hipótesis planteadas al inicio de la investigación. Como resultado al análisis, en las siguientes secciones se describen los resultados obtenidos que permiten validar el modelo conceptual multi-dominio para la representación del conocimiento aplicado en diferentes dominios, específicamente para esta investigación, aplicado en el dominio de la industria automotriz y el dominio de la industria de la innovación y el emprendimiento empresarial, para ambos dominios se utilizará la red social de confianza Twitter, como una fuente para la extracción de la información en los medios sociales en Internet.

6.2. Validación del modelo de representación.

Las hipótesis planteadas al inicio de la investigación proyectan la definición de un modelo para la representación del conocimiento a partir de la extracción de la información de los medios sociales. Por lo tanto, en la validación el modelo propuesto debe poder representar el conocimiento sobre los distintos productos, marcas, sus características y servicios ofertados en las redes sociales. Por otro lado, para la validación del modelo propuesto en esta investigación, se parte de dos casos de estudio que permite validar concretamente las hipótesis planteadas al inicio de la investigación. En este sentido, es necesario recordar cada una de las hipótesis y comentar al final del capítulo como se han comprobado cada una de ellas. Las hipótesis planteadas al inicio de la presente investigación son las siguientes:

H1: Mediante el diseño de un modelo conceptual multi-dominio es posible la caracterización de productos, medios sociales y usuarios, para gestionar el conocimiento y ser procesado por Sistemas Basados en Conocimiento.

- a) Mediante los medios sociales en la Web es posible caracterizar diferentes productos y perfiles de usuarios a través del uso de las Tecnologías Semánticas en la Web.
- b) Mediante la capacidad semántica y de soporte para contenido estructurado del modelo conceptual multi-dominio es posible recoger la publicación, edición e intercambio de información en los medios sociales.

Esta hipótesis desglosa, a su vez, dos hipótesis secundarias. Las cuales se han comprobado a partir del diseño de un nuevo modelo conceptual y, la definición de los conceptos propuestos en el Capítulo 4. Asimismo, a partir del modelo conceptual propuesto extraer la información de los medios sociales, a su vez, almacenarla en una ontología diseñada para esta investigación.

H2: Es posible ofrecer una solución escalable mediante un nuevo modelo conceptual genérico para recomendaciones en entornos multi-dominios y que sirva como una base para la operabilidad de Sistemas Basados en Conocimiento.

- a) Es posible la creación de un modelo conceptual genérico que sirva como base para Sistemas Basados en Conocimiento.
- b) Es posible la creación de una ontología del conocimiento orientada a la generación de recomendaciones personalizadas.

Esta hipótesis, al igual que la anterior desglosa dos hipótesis secundarias. Estas hipótesis se han comprobado a partir de la funcionalidad del marco de trabajo que integra una herramienta de recomendación. La generación de recomendaciones es basada en el modelo conceptual diseñado en esta investigación. Por lo tanto, la generación de las

recomendaciones mediante el marco computacional y, basadas en el modelo propuesto verifican estas hipótesis. El proceso se presenta en la Sección 6.4 del presente capítulo.

H3: Es posible diseñar una plataforma que integre el modelo conceptual multi-dominio basado en una solución sobre los contenidos en la Web, medios sociales y usuarios, mediante la anotación semántica y el contenido estructurado.

Esta hipótesis, se ha validado a partir de la funcionalidad del marco de trabajo basado en el modelo conceptual propuesto, asimismo integra una herramienta de recomendación para la solución de contenidos en la Web.

H4: Es posible la creación de un modelo conceptual genérico, adaptable a cualquier dominio de contenido estructurado y conocimiento semántico.

Esta hipótesis se ha comprobado a partir del diseño de un nuevo modelo conceptual multi-dominio descrito en el Capítulo 4 de la presente investigación.

Dentro de esta perspectiva, además de ratificar la comprobación de cada una de las hipótesis, la siguiente sección describe la validación por medio de la aplicación de un nuevo modelo conceptual multi-domino para recomendaciones mediante el filtrado de información semántica en los medios sociales. En la validación del modelo se ha representado el conocimiento de distintos dominios a partir de dos casos de estudio. En la siguiente sección se describe el primer caso de estudio presentado en esta investigación, el caso de estudio permite aplicar el modelo conceptual multi-dominio a un entorno real de trabajo.

6.3. Casos de estudio para la aplicación del modelo conceptual multi-dominio

Para llevar a cabo la aplicación del modelo se parte de un caso de estudio que utiliza la red social Twitter. Para el caso de estudio, se realiza un análisis sobre la aplicación del modelo conceptual multi-dominio propuesto en esta investigación. De esta forma, el modelo se puede aplicar a distintos dominios de contenido estructurado y semántico. Sin embargo, en esta sección se describe el primer análisis en un dominio concreto, el dominio de la industria automotriz. Un segundo análisis, permitirá trasladar al modelo conceptual propuesto a un segundo dominio. Por lo tanto, se aplicará mediante un segundo caso de estudio que será un análisis desarrollado en el dominio de la innovación y el emprendimiento empresarial, el segundo caso de estudio es presentado en la Sección 6.3.2.

Por lo tanto, en el modelo conceptual propuesto, se utiliza la red social Twitter como una herramienta para el análisis de los tweets, ya que como antes se ha mencionado, esta red social permite a los usuarios escribir mensajes de texto plano de corta longitud, además permite comentar sobre una variedad de temas y diferentes dominios que existen en la Web. Para este caso de estudio, se analizaron un total de 1.377.713 de tweets para un período comprendido entre el 25/06/2014 al 25/09/2014. Además de esto, para este caso de estudio se ha seleccionado la información relativa a las características asociadas con la industria automotriz, por lo que, mediante el seguimiento de un conjunto de marcas, un experto establece las características. Por consecuencia, la información es utilizada por un Sistema de Recomendación para la generación de recomendaciones hacia los usuarios externos.

6.3.1. Dominio en la industria automotriz

Como se ha comentado y documentado en los primeros capítulos, en la web, los usuarios hablan sobre una variedad de temas y productos, pero en esta investigación un tema relativo a la industria automotriz ha sido seleccionado para analizar el comportamiento en la aplicación del modelo conceptual propuesto en esta tesis. De tal forma que, los usuarios mencionan a través de una red social productos y sus características que mantienen una relación con las diferentes marcas del dominio de la industria automotriz.

Por lo tanto, para este caso de estudio, el experto en el dominio a través de la herramienta de modelado ha seleccionado las siguientes marcas: BMW, Audi, Fiat, Ford, Honda, Hyundai, Nissan, Peugeot y Renault. Estas marcas fueron elegidas para delimitar el análisis de contenido. Para el modelado de los productos, la información pertenece a una clase de tipo <<Categoría de productos>>, esta clase permite almacenar cualquier dominio de contenido estructurado y semántico para ser modelado, en este caso, la clase almacena los datos del dominio de la industria automotriz. Por consiguiente, la categoría se encuentra asociada con una clase de tipo <<Producto>>, la cual almacena los productos en el dominio de la industria automotriz, para este caso, se refiere a cada una de las marcas de productos en el dominio. En consecuencia, mediante el análisis de este contenido almacenado en las distintas clases definidas en el modelo propuesto, es posible identificar y caracterizar los productos que se refieren a las distintas marcas de automóviles, además de caracterizar sus contenidos de las marcas que mantienen una relación con las palabras empleadas por los usuarios a través de la cuenta de Twitter. Las palabras claves son definidas por el experto del dominio.

Con el fin de comprobar una de las hipótesis planteadas en la presente investigación, se realiza un experimento que permite analizar la ontología del dominio, sus relaciones y sus conceptos definidos en el modelo ontológico descrito en la Sección 4. Asimismo, la Figura 6:1 muestra un fragmento de la taxonomía de la ontología para el modelado en el dominio de la industria automotriz, en la figura se puede apreciar las distintas marcas de automóviles que pertenecen a la industria automotriz.

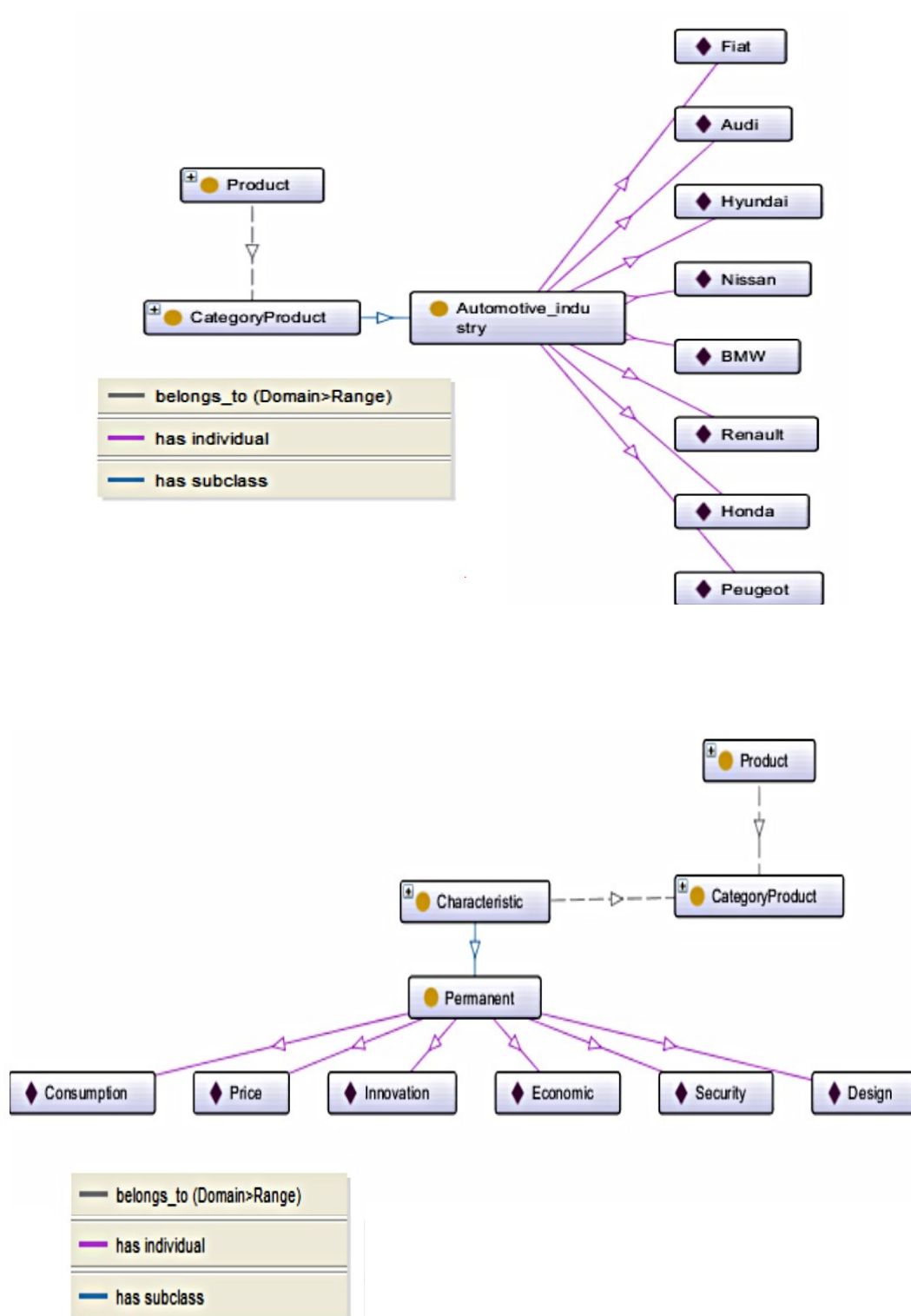


Figura 6:1 Fragmento de la taxonomía en el dominio de la industria automotriz.

Por otro lado, la Figura 6:2 representa la aplicación del modelo propuesto, sobre los principales conceptos del modelo formal aplicados al dominio de la industria automotriz y descritos en la Sección 4.5 de la presente tesis.

Por lo tanto, dichos conceptos y relaciones parten del denominado modelo formal presentado en la Sección 4.4. En este sentido, los conceptos son las distintas categorías definidas en las diferentes clases y conceptos de la ontología propuesta y descrita en el Capítulo 4. Asimismo, se representa un concepto de la marca del automóvil llamado <<Ford>>, que a su vez, dicha marca pertenece a una clase de tipo <<Producto>>. Así mismo, se representa el tipo de categoría para los conceptos de red social y una característica de seguridad.

La prueba para el análisis de la representación del conocimiento de estos conceptos se llevó a cabo utilizando el mismo período mencionado en la Sección 6.2, donde los usuarios mencionaron productos relacionados con las marcas de automóviles y características de seguridad.

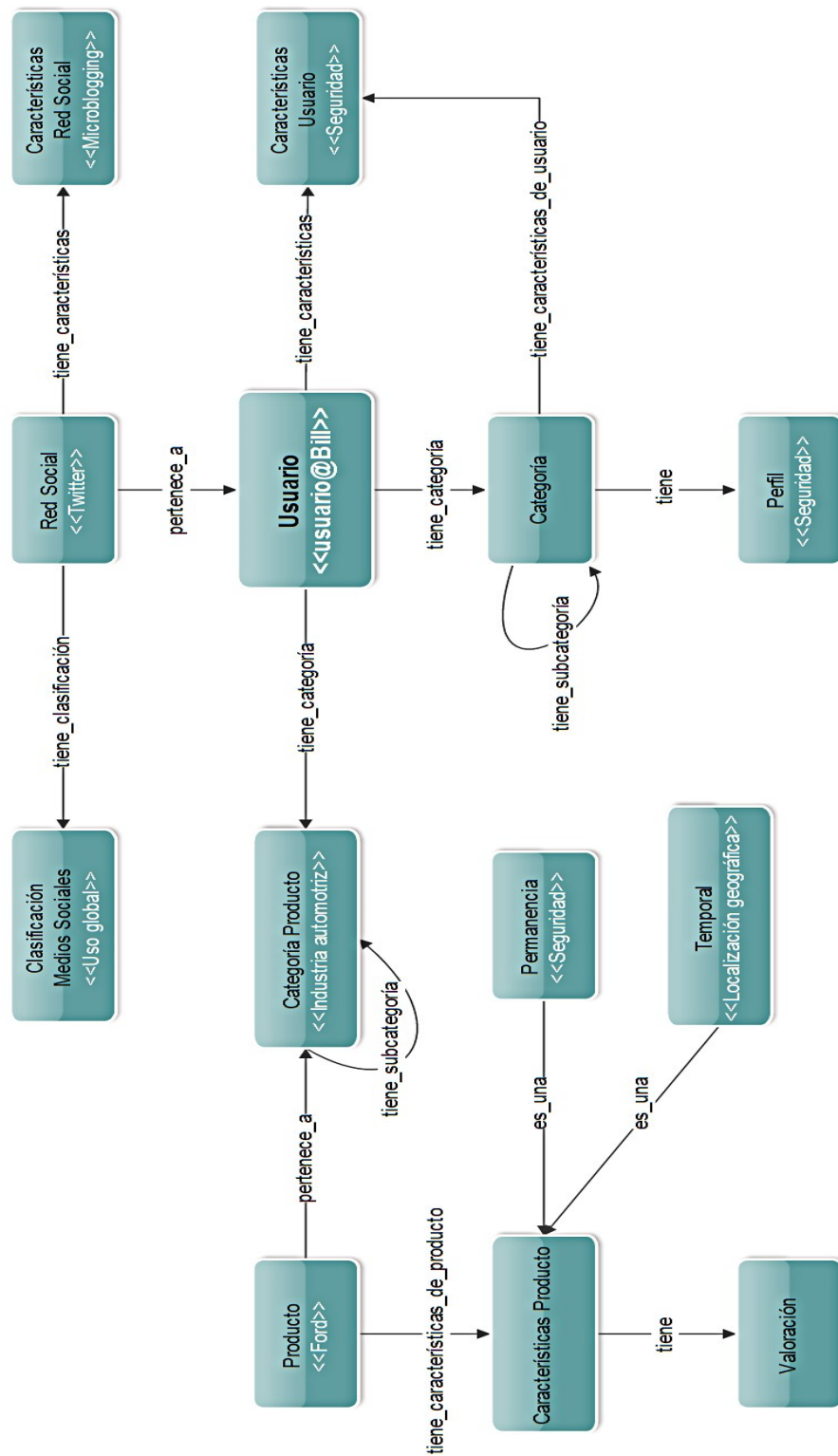


Figura 6:2 Conceptos y relaciones a partir del modelo formal aplicados al dominio de la industria automotriz.

Por lo tanto, un usuario tiene un tipo de producto que pertenece a un tipo de categoría, en este caso la categoría a la industria automotriz y un producto caracterizado que es denominado <<Ford>>. De manera que la categoría se encuentra relacionada con un producto caracterizado de la industria automotriz. De la misma forma, el usuario tiene características asociadas a una red social, que en este caso es la red social Twitter donde el usuario realiza menciones sobre las características del producto, para este caso menciona las característica de seguridad.

En consecuencia, es una característica relacionada con el tipo de concepto que a su vez, se encuentra relacionado con el usuario, un perfil que es conocido por la información mencionada en Twitter, la información es identificada y analizada por un experto. Sobre las características de un perfil, en esta investigación se caracterizan las menciones como una preferencia positiva y objetiva a través del usuario. Por ejemplo, en este caso el usuario habla de <<seguridad>>. Por lo tanto, las reglas del sistema caracterizan que un usuario le gusta la seguridad.

La Figura 6:3 muestra la evolución de las menciones de las distintas marcas a partir de la red social Twitter dadas por los tweets, de la que se deducen las relaciones de los conceptos descritos anteriormente (véase la Figura 6:2). Además, se muestra el número acumulativo de menciones para el análisis en el mismo período establecido.

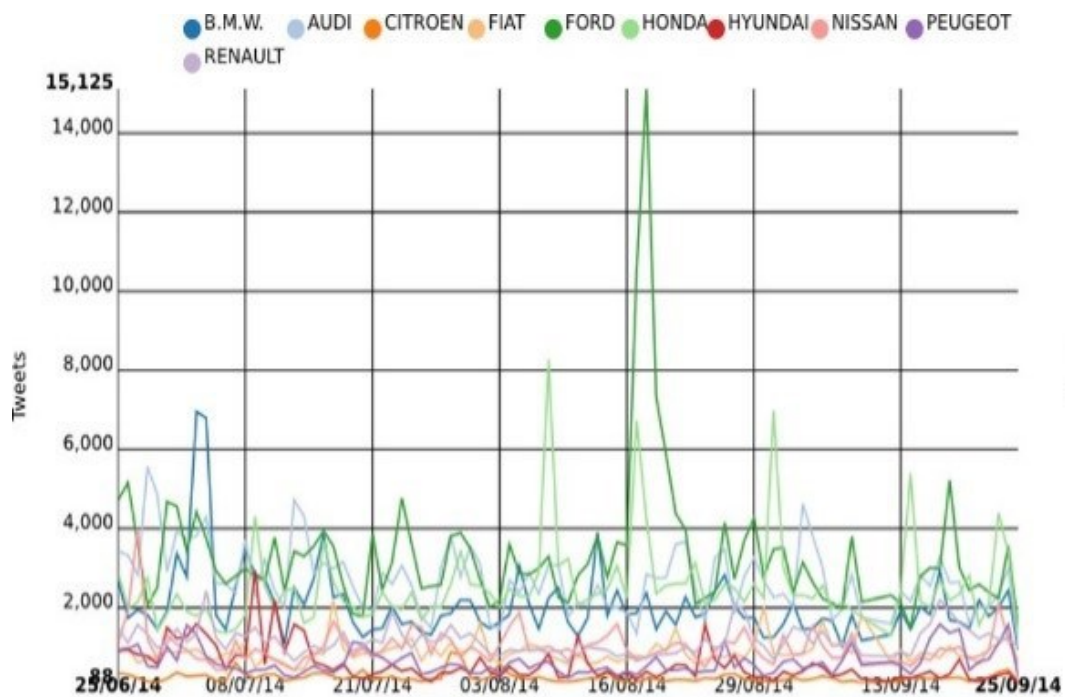


Figura 6:3 Gráfica sobre las menciones que realizan los usuarios hacia los productos del dominio de la industria automotriz en el rango establecido.

Los conceptos se asocian a los usuarios, los automóviles y, a su vez, con diferentes productos denominadas como marcas. Así mismo, a través de la extracción 1.377.713 tweets son analizados, con distintas marcas de automóviles que son mencionados por los usuarios.

La Figura 6:4 muestra el número acumulativo de menciones en el mismo período establecido. De manera que es posible identificar y caracterizar las distintas marcas que mencionan los usuarios a través de la red social.

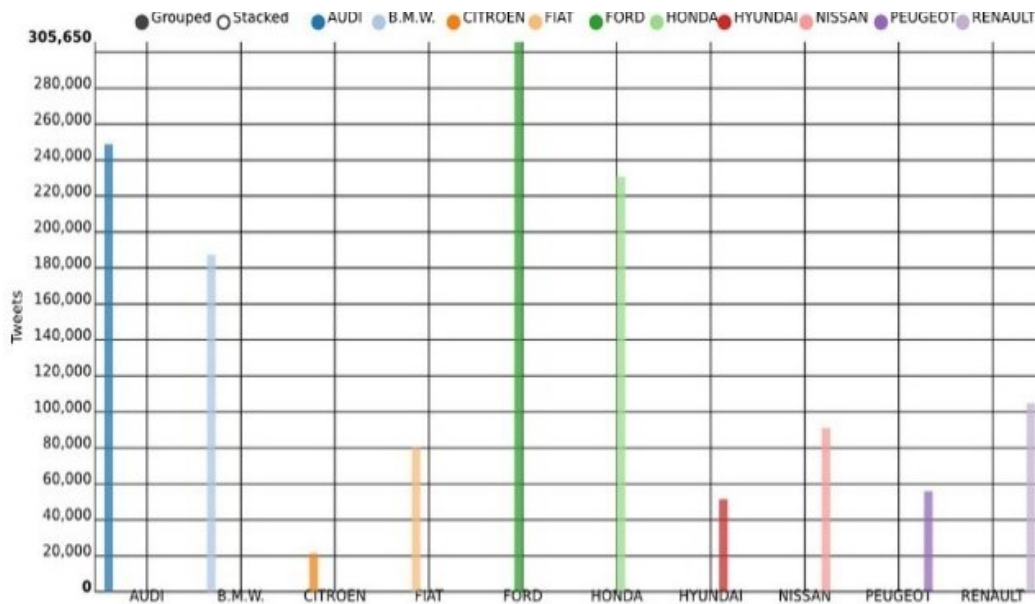


Figura 6:4 Gráfica de los productos clasificados por marcas.

Los conceptos propuesto en el modelo conceptual, permiten asociar a usuarios, automóviles y, a su vez, los diferentes productos denominados en este estudio como marcas. Así mismo, a través de la extracción, se analizan 1.377.713 tweets con distintas marcas de automóviles que son mencionados por los usuarios.

Con la información que se extrae es posible identificar y caracterizar las distintas marcas que se han mencionado en la red social. Por lo tanto, las relaciones del modelo formal apoyan los resultados del análisis. Como consecuencia, las diferentes marcas de automóviles que han sido mencionados por los usuarios de Twitter pueden ser caracterizadas. Además, la información caracterizada puede ser utilizada a modo de preferencias para futuras recomendaciones a partir de un SR integrado en el marco computacional de trabajo propuesto en el Capítulo 5.

Por consiguiente, en el análisis que arrojan las gráficas de la Figura 6:3 y la Figura 6:4, se puede observar que el producto <<Ford>> es la marca de automóviles que los usuarios han mencionado con mayor constancia, con 305.650 tweets.

Por otro lado, la Tabla 6.1 describe los conceptos que apoyan el modelo formal descrito en el Capítulo 4, a su vez, mantienen relación con los conceptos de la Figura 6:2. Para aplicar el modelo a dos casos de estudio, dichos conceptos son utilizados para especificar los valores obtenidos para este caso de estudio.

Tabla 6:1:

Descripción de los conceptos del modelo formal para el caso de uso aplicado al dominio de la industria automotriz.

CONCEPTO	VALOR
Usuario	user@Bill_(nombre_ficticio)
Categoría de producto	Industria automotriz
Producto	Ford
Características de producto	Permanencia / Temporal
Red social	Twitter
Clasificación de red social	Uso global
Características de red social	Microblogging
Categoría	Perfil
Características de usuario	Seguridad

Del mismo modo, en la Figura 6:5, se muestran las relaciones y características entre los conceptos que pertenecen a un tipo de clase definida en la ontología del sistema. Por lo tanto, para este análisis se describen conceptos relacionados con las características de un producto llamado <<Audi>>. Además, el producto pertenece a la misma categoría de tipo automóvil. Además, el producto tiene una característica que se llama <<Seguridad>>, en el que además se trata de una característica de interés para el usuario, ya que las reglas del sistema caracterizan que un usuario le gusta la <<Seguridad>>.

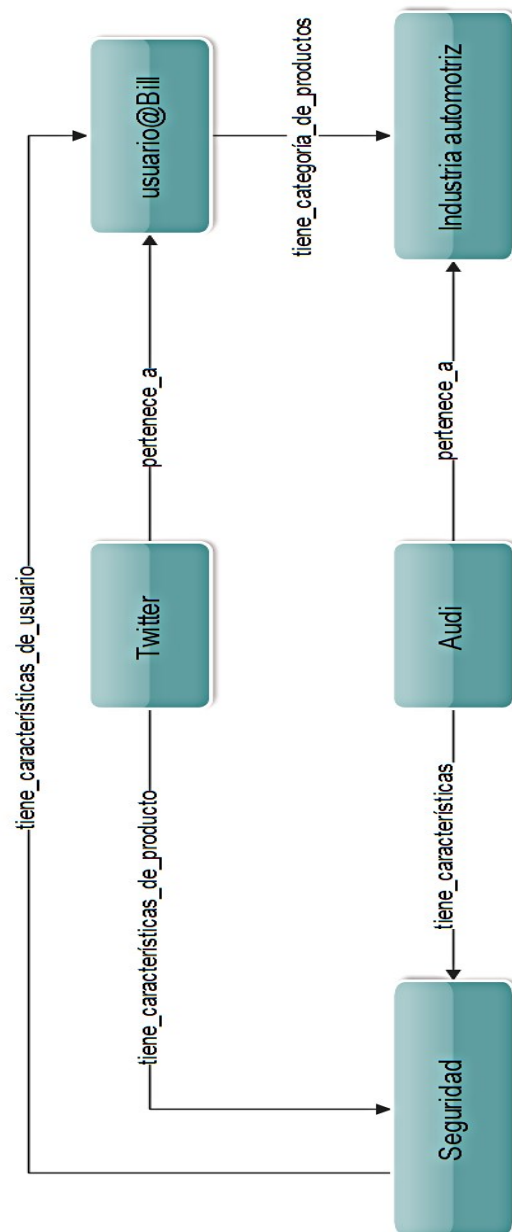


Figura 6:5 Relaciones y características.

Por lo tanto, para este análisis los usuarios mencionan la palabra seguridad en sus tweets. Además, el experto expresa que tienen un perfil corporativo, ya que mencionan un producto de la marca <<Audi>>. En la gráfica de la Figura 6:6, se muestra un análisis de los usuarios que mantiene un perfil corporativo ya que mencionan la marca <<Audi>>

y, asimismo la gráfica muestra el número de usuarios que mencionan la característica de <<Seguridad>>.

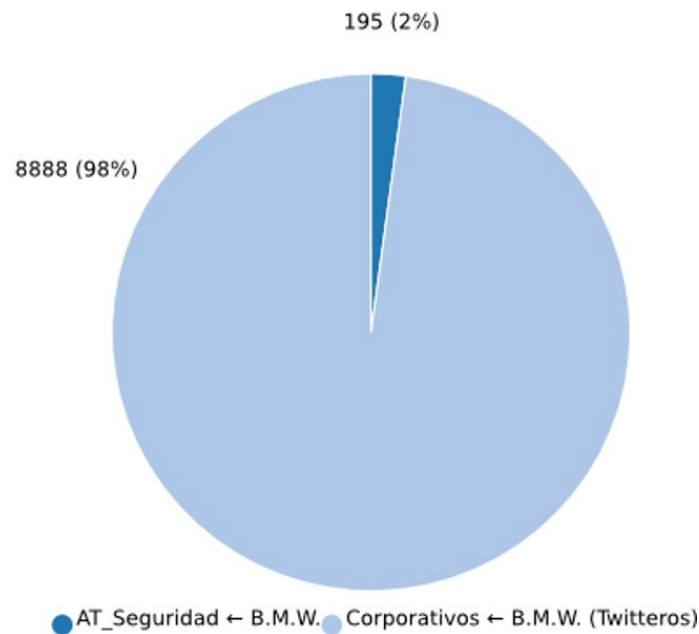


Figura 6:6 Análisis que menciona relaciones de características.

Para este experimento que soporta la validación del modelo propuesto, se analizaron 9,0830 tweets, de los cuales 195.000 corresponden a un dos por ciento que mencionan la característica de <<Seguridad>>. Por lo tanto y llegado a este punto del análisis, se reafirman la comprobación de las hipótesis formuladas al inicio de la presente tesis doctoral, ya que es posible la caracterización de perfiles y productos a partir de las menciones que realizan los usuarios en la red social.

Además, tal y como se ha comentado con anterioridad las reglas del sistema caracterizan las menciones como una preferencia positiva y objetiva, de manera que a un dos por ciento de los usuarios analizados en este experimento se afirma que les interesa la seguridad. La información es almacenada en la base de conocimiento y es utilizada en futuras recomendaciones a partir de un sistema de recomendación.

En este mismo sentido, los conceptos establecen una relación entre los intereses de los usuarios en base a las entradas realizadas con dichas características. En consecuencia, se comprueban las primeras dos hipótesis, ya que el experimento permite por un lado, la extracción de la información y por otro lado la caracterización de un producto, un usuario y una red social. Por lo tanto, el modelo conceptual sirve para representar y generar contenido estructurado y semántico, además el contenido sirve para recomendaciones y como una base para la operabilidad de Sistemas Basados en Conocimiento.

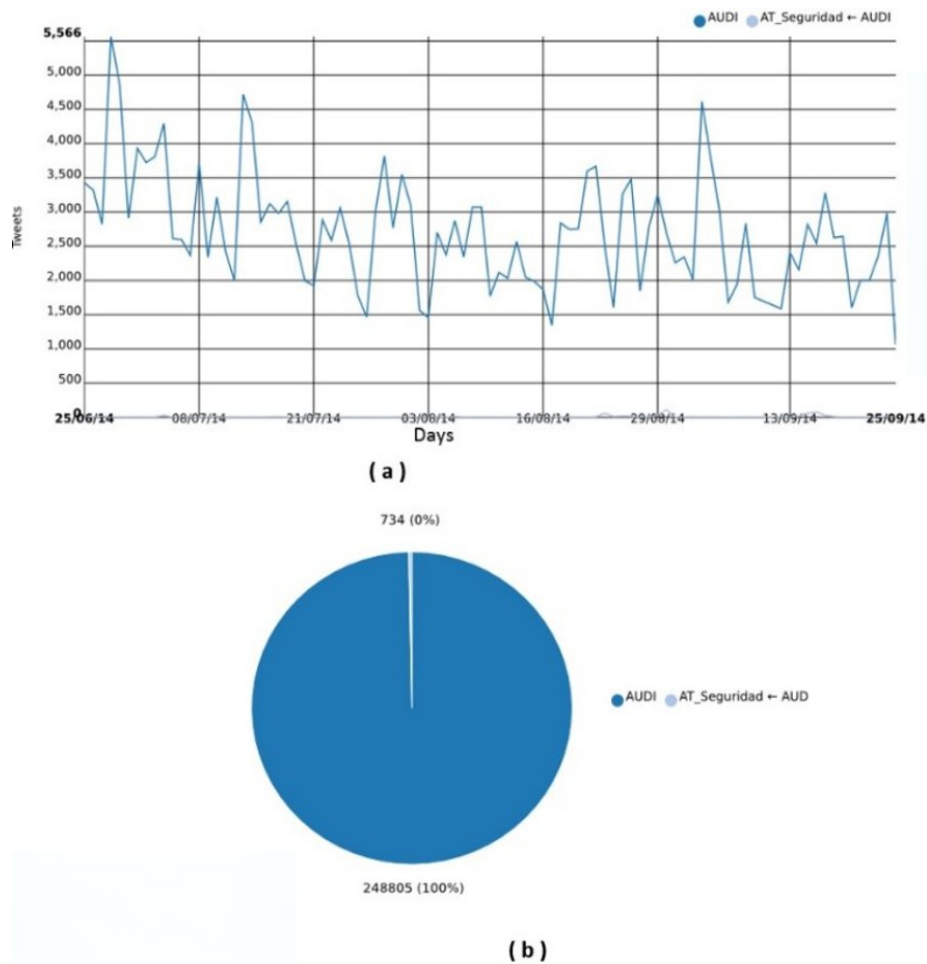


Figura 6:7 (a) Tweets que mencionaba marca Audi, (b) Tweets que mencionan característica de seguridad

La Figura 6:7 permite visualizar, por un lado, una gráfica sobre los usuarios que a través de los tweets mencionan a una marca, para este caso, la marca de coche <<Audi>>. Por otro lado, una gráfica que permite visualizar los tweets que los usuarios mencionan para hacer referencia a una característica de la marca, en este caso la característica es <<Seguridad>>. Los resultados del análisis respaldan a los conceptos mostrados en la Figura 6:5. Además de esto, es compatible con la relación que se establecen entre los conceptos de las clases características, productos, red social, categorías y usuarios del modelo formal descrito en el Capítulo 4.

Por consiguiente, se extraen 248.805 tweets que mencionan un producto llamado <<Audi>>, de los cuales 734,00 mencionan la característica de <<Seguridad>>. La información para este análisis sirve para identificar y demostrar que los usuarios hablan de la característica de seguridad.

En este sentido, además no solo para una marca en específico sino para distintas marcas de productos relacionados con el dominio de la industria automotriz. Por lo tanto, es posible llevar a cabo una caracterización del contenido sobre las marcas de automóviles. Además, es clara la importancia de la identificación de las preferencias de los usuarios con respecto a aquellas características que mencionan a través de la red, la información podría ser utilizada para futuras recomendaciones a partir de un SR.

La gráfica (b) de la Figura 6:7, muestra la extracción de información relativa a la marca <<Audi>> y sus características de <<Seguridad>>. De hecho, con esta información permite conocer el interés del usuario sobre un producto a ser conocido; esto es con el fin de analizar el contenido extraído de la red social Twitter que también nos ayuda a conocer las preferencias de los usuarios. Además, sobre la base del modelo semántico, la información del producto, en este caso la marca <<Audi>> y sus propias características, como lo es la característica de seguridad, se visualizan. Por lo tanto, esta característica nos permite identificar los intereses del usuario para la validación de un producto.

Por consiguiente, en la Figura 6:8, se muestran las relaciones de alto nivel de los conceptos que permiten validar el caso de estudio a partir del modelo conceptual

propuesto. Estos conceptos son la llave que permiten que la información obtenida a partir de datos de la Web (ver Sección 5.3) puedan ser modelados y validados.

Por otra parte, también se representan las relaciones de los conceptos para un modelado entre un usuario que tiene la red social Twitter y el nombre del producto. Este concepto pertenece a una categoría de tipo automóvil, el producto ahora mantiene otras características que mencionan los usuarios, para este caso las características de precio y consumo.

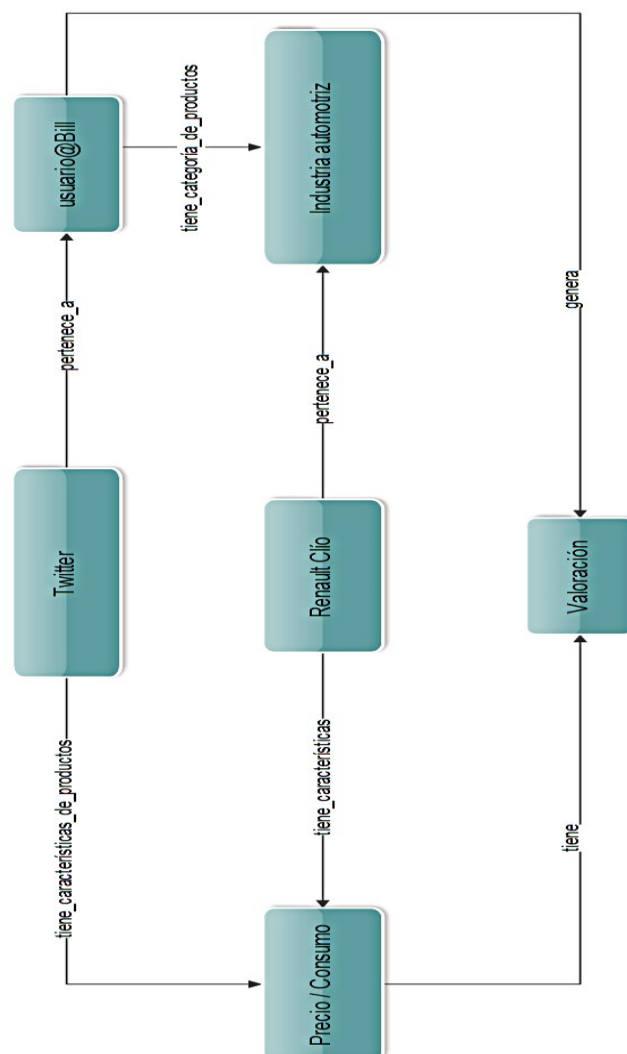


Figura 6:8 Relaciones de alto nivel para la validación a partir de un caso de estudio.

En este sentido, se realizó otra prueba utilizando el mismo período ya mencionado en la Sección 6.3.1. Por lo tanto, para este caso de estudio 2.778 tweets han sido extraídos para realizar un análisis que permita validar las relaciones entre los conceptos de alto nivel propuestos a partir del modelo conceptual.

Como resultado en esta prueba, el experimento permite analizar los tweets de los usuarios que mencionan un producto denominado <<Renault Clio>>, el producto mantiene las características de <<Precio y consumo>>.

El usuario también tiene una relación con la valoración de estas características. Con el fin de satisfacer los intereses de los usuarios, y de modo que en el futuro la información servirá para hacer recomendaciones a través de un sistema de recomendación, la información se almacena en la ontología del sistema. Varios métodos pueden ser aplicados a la valoración. Por ejemplo, se puede basar en los resultados de sentimiento que permiten a las opiniones de los usuarios que se conocen en relación con su clasificación de productos a través de la positividad y la negatividad de las palabras por medio de vocabulario como SentiWordNet. Pero en esta investigación se ha optado por la positividad y objetividad de las menciones que el usuario realiza a través de sus tweets. De manera que el sistema utiliza reglas que caracterizan positiva y objetivamente las menciones.

Por lo tanto, esta información sirve para caracterizar semánticamente las marcas y productos del dominio. Además es posible incorporar nuevos conceptos en la base de conocimiento que puedan ser utilizados por la ontología del sistema. Por ejemplo los hashtags de los tweets. Se refleja en la relación entre el usuario, el rating y los conceptos de precio / consumo.

Por lo tanto, en esta prueba 104.847 tweets son analizados donde los usuarios han mencionado un producto de la marca <<Renault Clio>> y las características del producto que son <<Precio / consumo>>, la prueba se realizó en el mismo período de tiempo establecido en la Sección 6.3.1. La Figura 6:9, muestra los resultados del análisis del producto y característica, la Figura 6:10 muestra los resultados obtenidos a través de las palabras claves seleccionadas y mencionadas por los usuarios.

Finalmente, la Figura 6:11 muestra el número de tweets que mencionan el producto y la característica, para este caso la marca <<Renault>> como un producto modelado en el dominio automotriz y las características de <<Precio / Consumo>> para el mismo producto modelado. La información es caracterizada en la base de conocimientos.

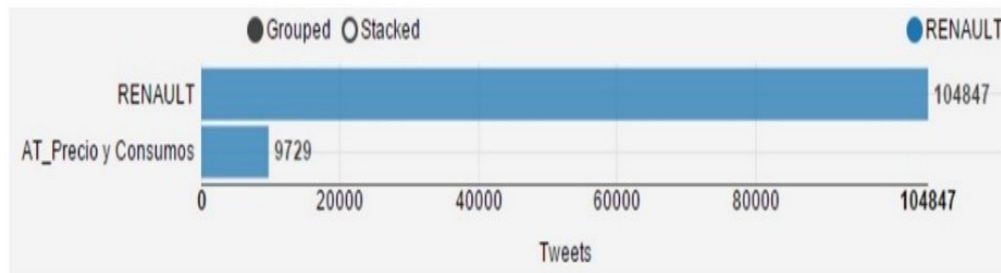


Figura 6:11 Número de menciones para un producto y sus características.

En un segundo experimento para el mismo caso de estudio, se analizaron 374.951 tweets. Para esta prueba, los usuarios mencionan productos sobre tres marcas de automóviles: <<Audi>> con 248,805.00 tweets, <<Renault>> con 104,847.00 tweets y <<Citroën>> con 21,299.00 tweets. Estas marcas pertenecen a la clase de tipo <<Categoría de producto>>. La Figura 6:12, muestra un gráfico donde se puede visualizar y analizar a los usuarios que mencionan los productos sobre las marcas de automóviles. Además, se muestra el porcentaje de los tweets para cada una de las marcas que son mencionadas en la red social Twitter.

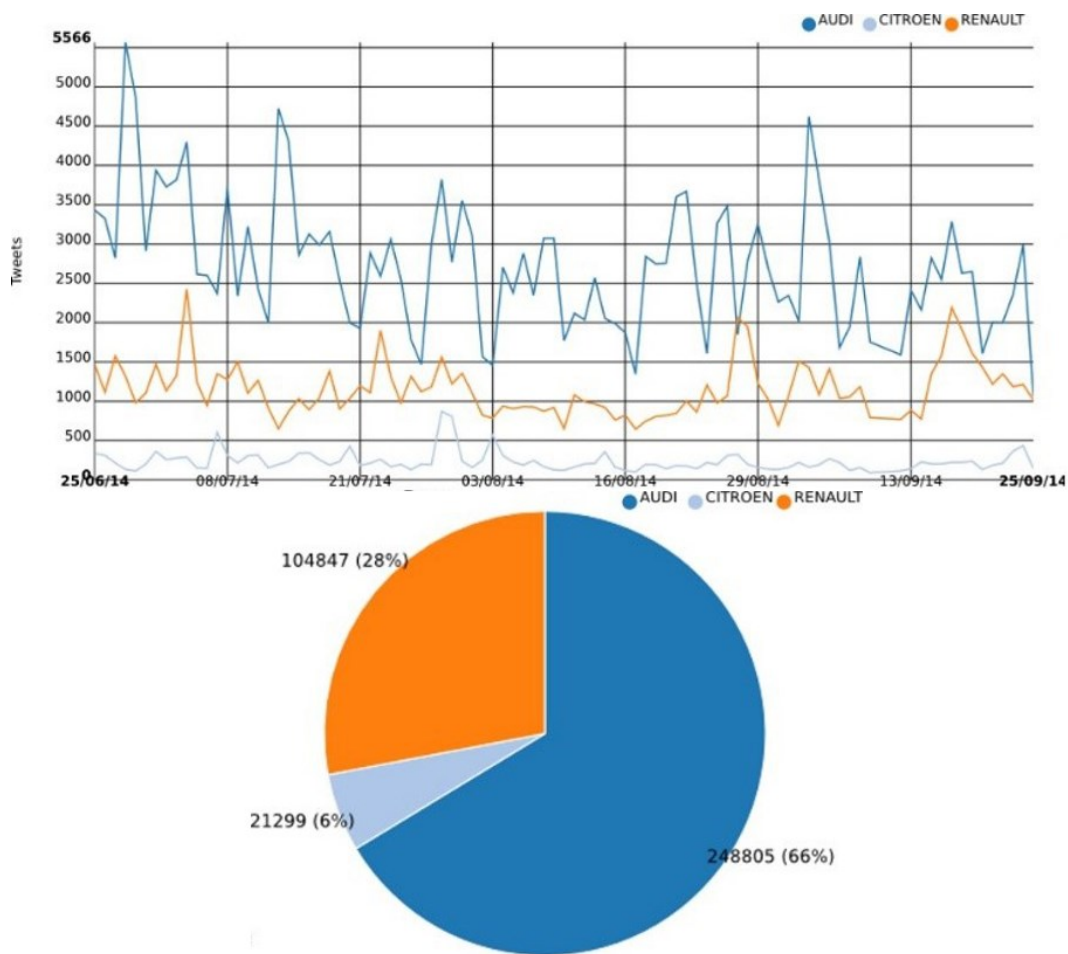


Figura 6:12 Análisis de Tweets que mencionan productos de marca Audi, Citroën y Renault.

Por otra parte, también se mencionan tres tipos de características. Por otro lado, la Figura 6:13 muestra a los usuarios que mencionan las características de <<Precio y consumo>> con 116 tweets; <<Diseño>> con 2465 tweets; <<Deportividad>> con 2314 tweets y <<Ecología>> con 2554 tweets. Asimismo, la Figura 6:14 muestra otro ejemplo para un análisis añadiendo una nueva característica, la de <<Tecnología>> con 5.453 menciones. Además de mantener las características de <<Diseño>> con 2.465 menciones y <<Precio / consumo>> con 116 menciones de los usuarios. Además, se muestra el porcentaje de tweets sobre las menciones que los usuarios realizan para cada característica y, que además vinculan la relación con sus intereses de los usuarios.

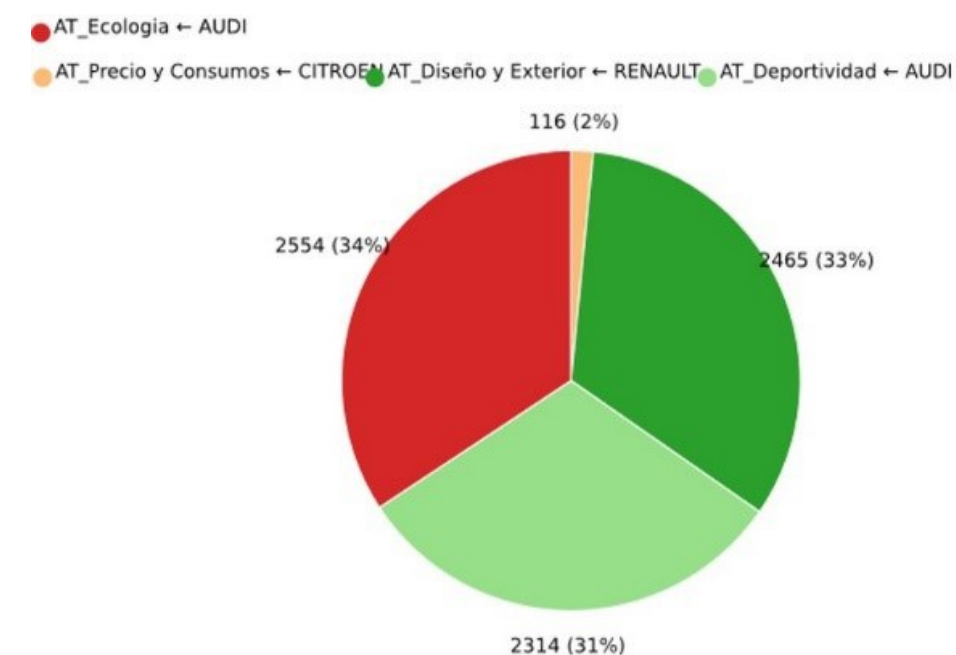


Figura 6:13 Características del dominio en la industria automotriz.

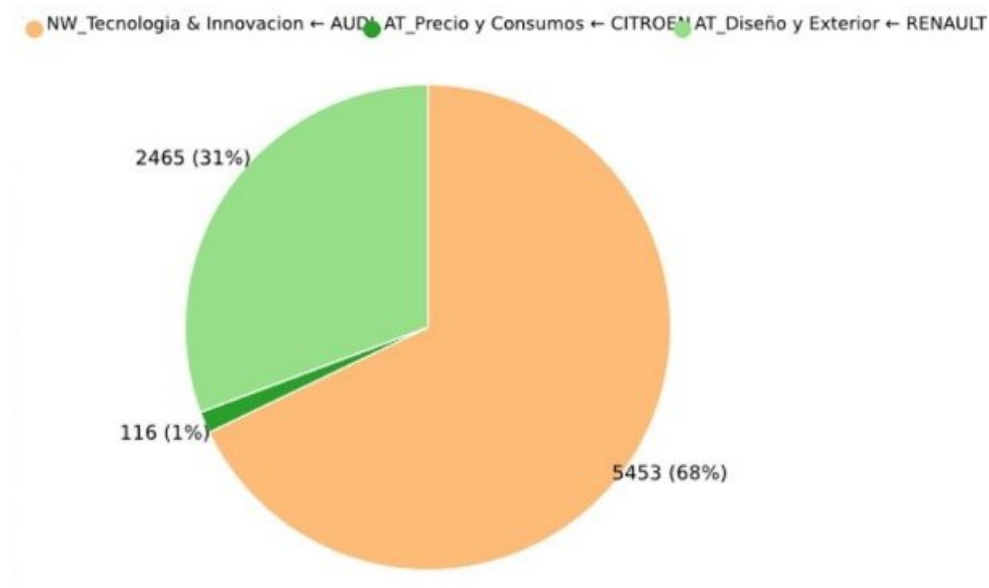


Figura 6:14 Tweets que mencionan diferentes características de productos

Por lo tanto, con estas pruebas, se ha comprobado que el modelo conceptual puede caracterizar un contenido estructurado sobre los distintos productos, marcas, sus características y servicios ofertados en las redes sociales, a su vez, puede modelar y gestionar el conocimiento de diferentes perfiles de usuarios, productos y medios sociales caracterizados en múltiples dominios.

No obstante, a continuación se describe otro experimento utilizado para el mismo dominio que permite reiterar el objetivo principal y la validación de las hipótesis del modelo conceptual.

En este experimento se realiza un análisis sobre las características de un producto en el dominio de la automoción. El producto es llamado <<BMW>> con características establecidas, tales como, <<Seguridad>>, <<Precio/consumo>> y <<Potencia>>. El perfil de usuario se establece a partir de las menciones sobre las características del producto que son expuestas por el mismo usuario a través de las redes sociales. La Figura 6:15, muestra las relaciones entre los conceptos del dominio a partir del modelo conceptual. Además de las características de <<Usuario>> y <<Producto>> a partir de la información extraída de la red social Twitter. El marco computacional utiliza la información para generar recomendaciones a partir de las menciones que los usuarios realizan.

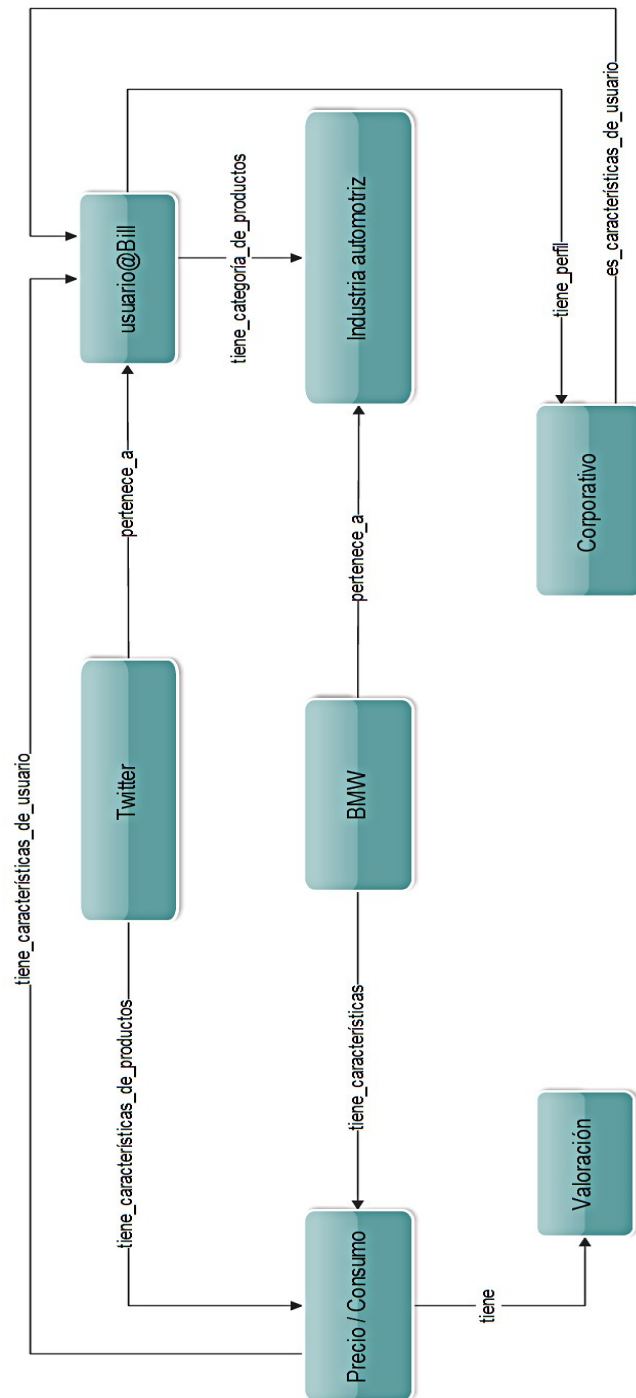


Figura 6:15 Relaciones entre los conceptos del dominio automotriz

En este experimento, para conocer las características del producto, se realiza un análisis donde los usuarios mencionan la palabra <<Precio/consumo>> a través de sus tweets.

Además, el experto establece que es un perfil corporativo, ya que menciona un producto de alta gama sobre la marca <<BMW>>. Por lo tanto, en la Figura 6:16 se muestra el análisis estadístico sobre usuarios con un perfil corporativo que han mencionado en sus tweets la marca <<BMW>>, además mencionan las características de <<precio/consumo>>.

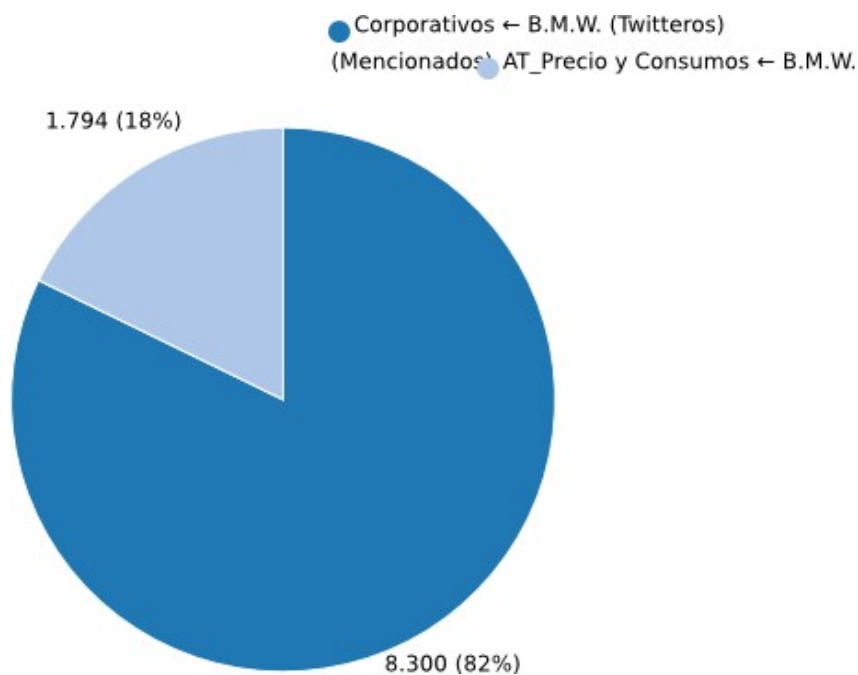


Figura 6:16 Análisis de características y perfiles

En esta prueba, se analizan 10,094 tweets, de los cuales 1,794 tweets corresponden a un dieciocho por ciento que mencionan características de <<Precio/consumo>>. El marco computación de trabajo establece un rasgo de interés en dicho producto y característica, por lo tanto la información es utilizada por el sistema para generar recomendaciones personalizadas. Además, la relación del usuario con una red social permite conocer la información para la caracterización de un producto y, al mismo tiempo, la mención de las características del producto por medio de un usuario, de manera

que la información permite establecer un tipo de perfil basado en un experto en el dominio. Los conceptos establecen una relación entre los intereses de los usuarios en base a las entradas realizadas con dichas características del producto. Durante el periodo establecido. En esta prueba se han analizado los tweets que se relacionan con las marcas y las características de la industria automotriz.

En este mismo sentido, una segunda prueba para el mismo caso de estudio permite analizar 334,36.00 tweets. En esta prueba, los usuarios mencionan tres tipos de marcas de automóviles: <<BMW>> con 187,330.00 tweets, <<Nissan>> con 90.974.00 tweets y <<Peugeot>> con 56,059.00 tweets. Estas marcas pertenecen a una categoría de producto denominada <<Automóvil>>, que es un concepto definido en la ontología del sistema. La Figura 6:17, muestra un gráfico del análisis donde los usuarios mencionan las tres marcas durante el período establecido. Además, se visualiza el porcentaje de los tweets para cada una de las marcas que son mencionados por los usuarios.

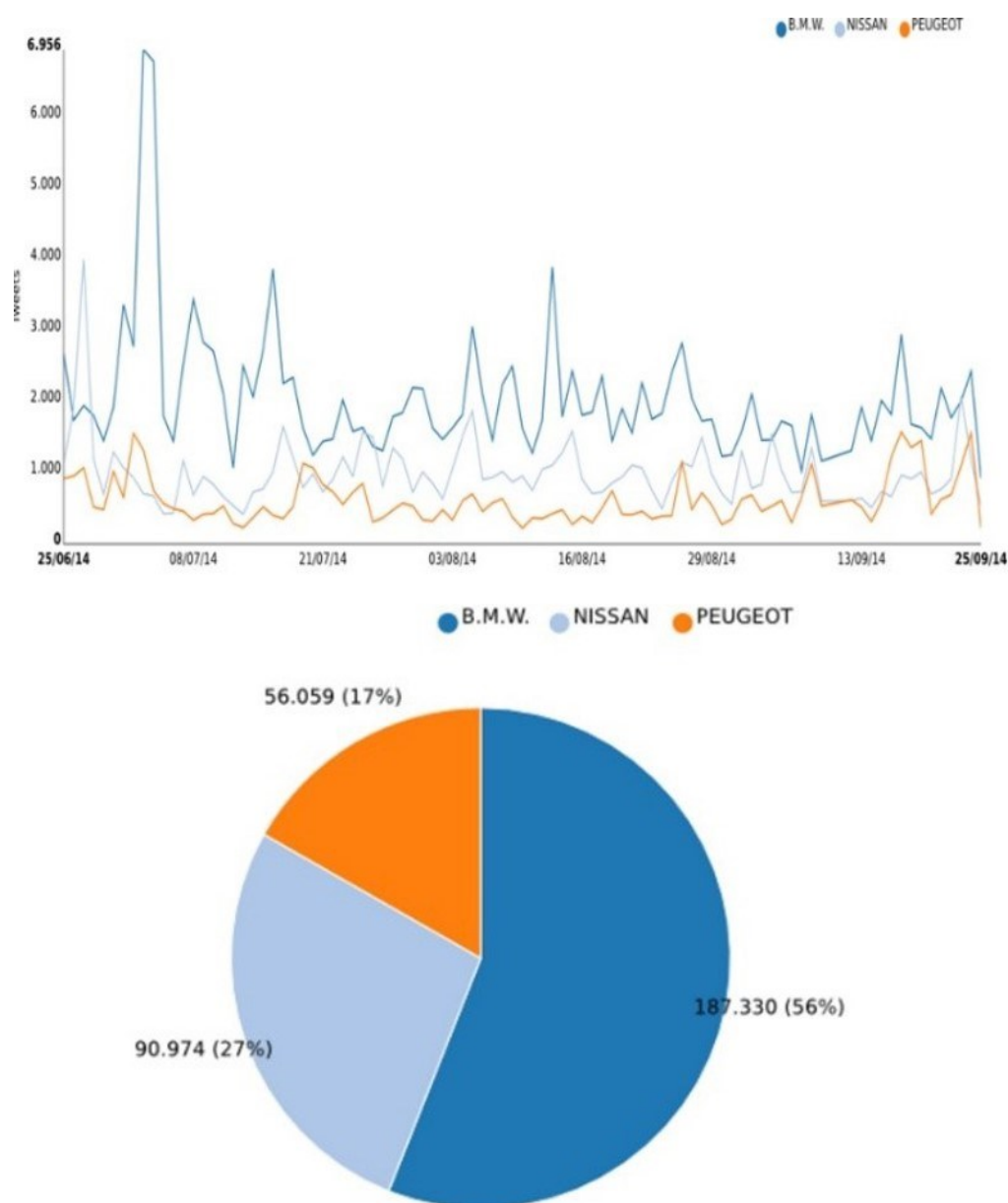


Figura 6:17 Usuarios que mencionan distintos productos a través de la red social Twitter

Por otra parte, los usuarios también mencionan distintas características para un producto. Por lo tanto, en esta prueba el experto establece diferentes tipos de características. En la Figura 6:18, se muestran las menciones que realizan los usuarios para productos sobre marcas de automóviles y características, tales como, <<Precio/consumo>> con 983 tweets; <<Diseño>> con 707 tweets; <<Deportividad>> con

2,314 tweets, <<Motor>> con 2,039 tweets, <<Ecología>> con 85 tweets, <<Seguridad>> con 1,399 tweets, <<Patrocinio>> con 4,360 tweets y finalmente la característica de <<Competencia>> con 18,078 tweets mencionadas por los usuarios.

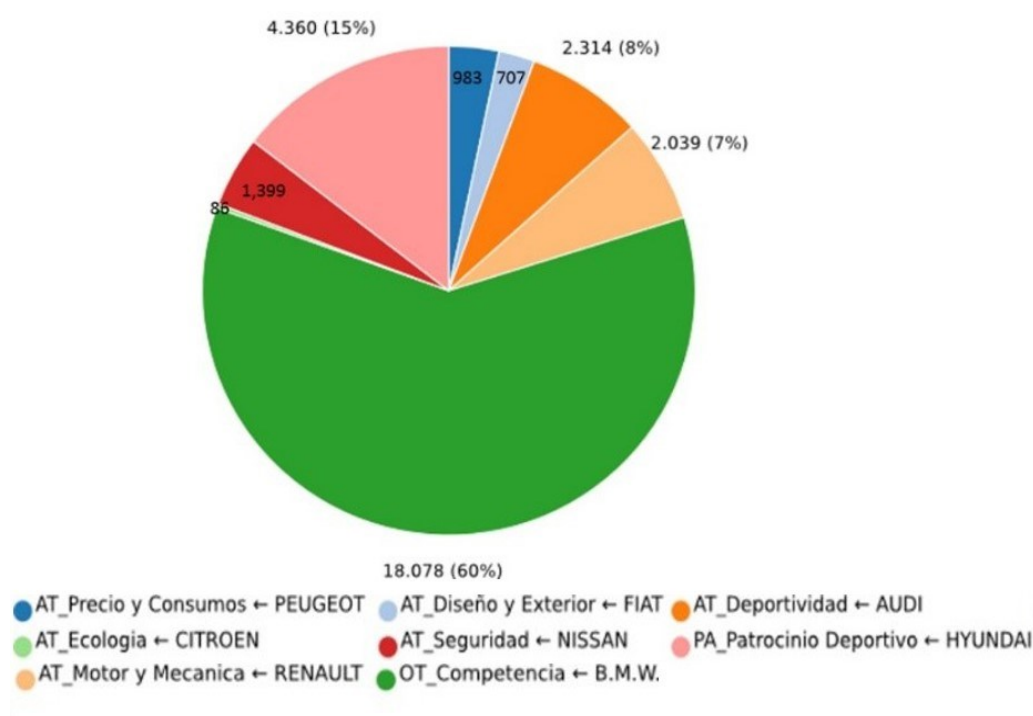


Figura 6:18 Usuarios que mencionan distintas características a través de la red social Twitter.

La clasificación de contenidos realizados en el marco computacional de trabajo sirve para la caracterización de contenidos, a partir de dicha información se generan recomendaciones personalizadas por medio de un Sistema de Recomendación hacia los usuarios sobre los productos almacenados en los medios de comunicación social. Además, el marco computación de trabajo establece perfiles de usuario a partir de las palabras claves que son mencionadas en los tweets, de manera que las características de los productos son de interés para un usuario

En esta sección se ha presentado un análisis del modelo propuesto aplicado al dominio de la industria automotriz. El análisis permite validar el modelo conceptual y las

hipótesis planteadas al inicio de la tesis doctoral. Además, a partir de las implementaciones el modelo ha sido evaluado en entorno computacional real. En la siguiente sección se describe un análisis del modelo propuesto para un dominio distinto.

6.3.2. Dominio en la industria de la innovación y el emprendimiento empresarial

Con el fin de reiterar la validación del modelo propuesto, se realiza un experimento que permite validar las relaciones y conceptos del modelo descritos en la Sección 4.2 para un dominio distinto. En este análisis se presenta un experimento relativo a la información sobre las características asociadas en el dominio de la iniciativa empresarial sobre la innovación y el emprendimiento. Para este análisis se utiliza el mismo periodo establecido para la extracción del contenido, además se seleccionaron los siguientes productos: la iniciativa empresarial cultural, la iniciativa empresarial digital, la iniciativa empresarial de la empresa, la iniciativa empresarial de turismo, la iniciativa empresarial tecnológica, iniciativa del emprendimiento energético, mujeres con iniciativa empresarial y el emprendimiento joven social. Estos productos fueron elegidos para delimitar el análisis de contenido, ya que pertenecen a una categoría de productos de la iniciativa empresarial, la categoría se encuentra asociada a un producto en el dominio de la innovación y el emprendimiento empresarial denominados. Por lo tanto, mediante el análisis de este contenido es posible identificar las características de los productos basados en las palabras claves que son mencionadas por los usuarios a través de la red social Twitter.

La Figura 6:19, muestra el modelo de los conceptos de innovación y emprendimiento, que a su vez representan una clase en la ontología del modelo, para este caso la clase <<Categoría de producto>>. Además, representa un concepto llamado <<Turismo>> que representa un tipo de producto de la marca, así mismo, un concepto que representa al tipo de categoría de la red social y una característica definida <<Innovación>>. Los usuarios mencionaron productos relacionados con el emprendimiento y características de la innovación. Por lo tanto, un usuario tiene una categoría de producto de la iniciativa empresarial, que en este caso sería <<Turismo>>, el

concepto se encuentra asociado con el dominio de la iniciativa empresarial. De la misma manera, el usuario tiene características de una red social donde se menciona la innovación, que en este caso es la red social Twitter. Además, una característica relacionada con el concepto de usuario es un perfil conocido por la información almacenada en Twitter, la información es identificada y analizada por un experto en el dominio.

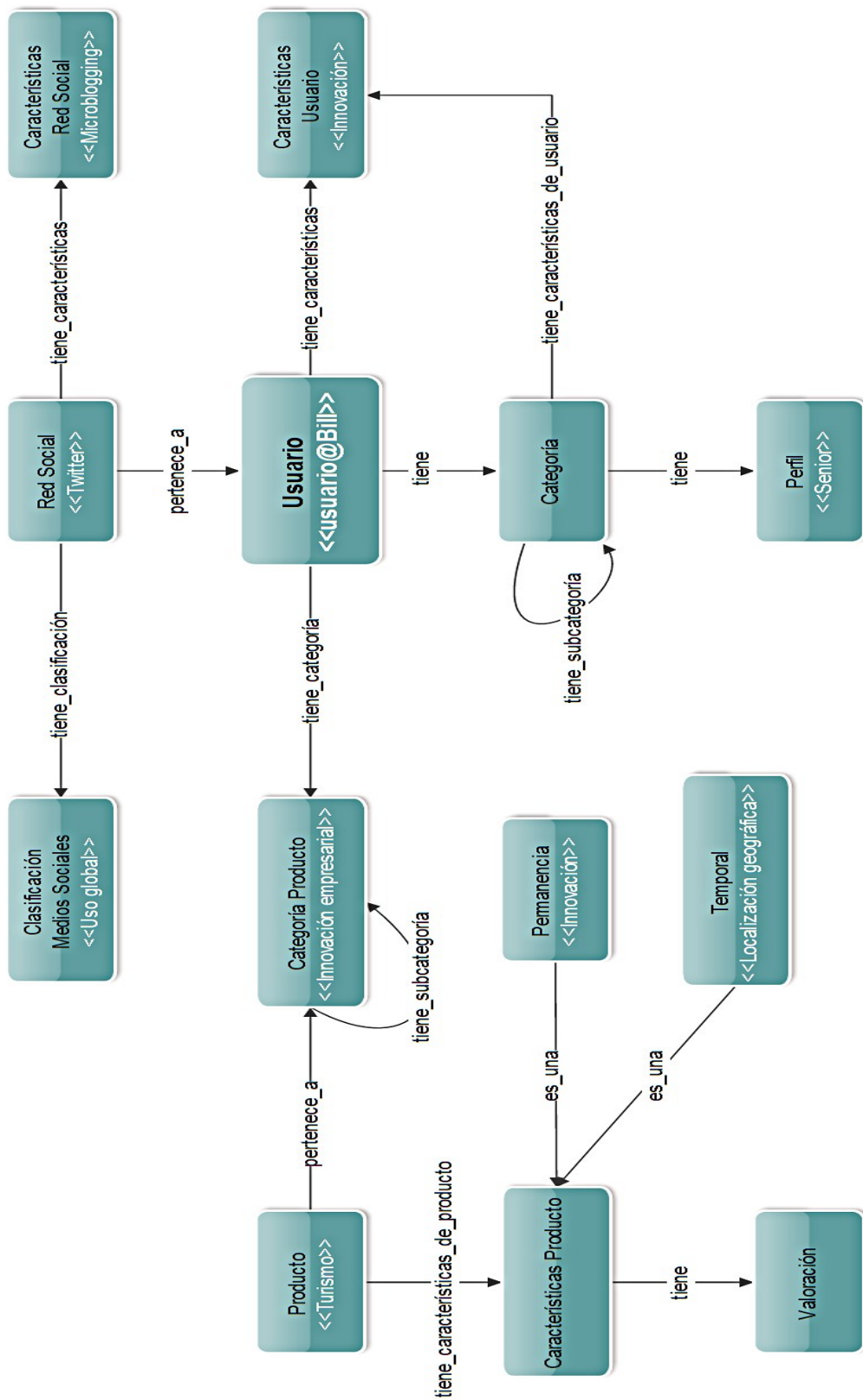


Figura 6:19 Conceptos y relaciones a partir del modelo formal aplicados al dominio de la industria de la Innovación y el emprendimiento empresarial.

La Figura 6:20, muestra la evolución de las menciones sobre las distintas marcas y productos en Twitter dadas por los usuarios y, aplicadas al dominio de la industria de la innovación y el emprendimiento empresarial, de la que se deducen las relaciones de los conceptos descritos anteriormente (véase la Figura 6:19). Además de visualizar el periodo para la extracción de los datos. Los conceptos se asocian a los usuarios, el emprendimiento empresarial y, a su vez, con diferentes productos. Por lo tanto, a través de dicha red y en relación con los diferentes productos de la iniciativa empresarial que se mencionan por los usuarios es posible realizar un análisis del contenido.

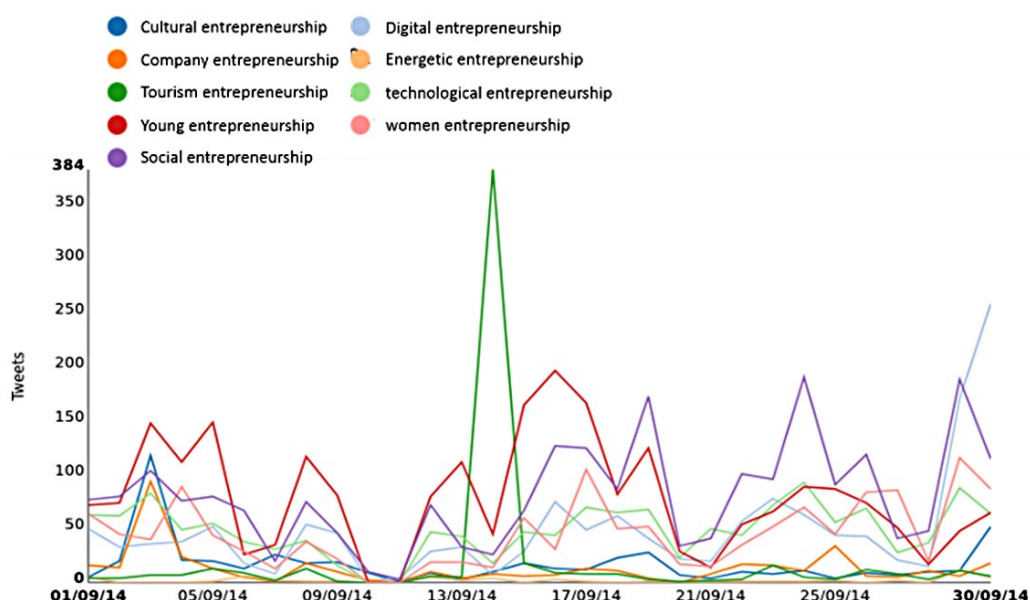


Figura 6:20 Evolución de las menciones de los usuarios sobre productos para el dominio de la industria de la Innovación y el emprendimiento empresarial.

La Figura 6:21, muestra el número acumulado de menciones en el mismo período establecido, donde es posible ver los diferentes productos que se han mencionado en la red social. Por lo tanto, las relaciones del modelo formal apoyan los resultados del análisis, así como los diferentes productos de emprendimiento que han sido mencionados por los usuarios de Twitter pueden ser identificados; Además, la información se almacena en la

base de conocimiento y sirve a modo de preferencias para futuras recomendaciones a partir de un SR.

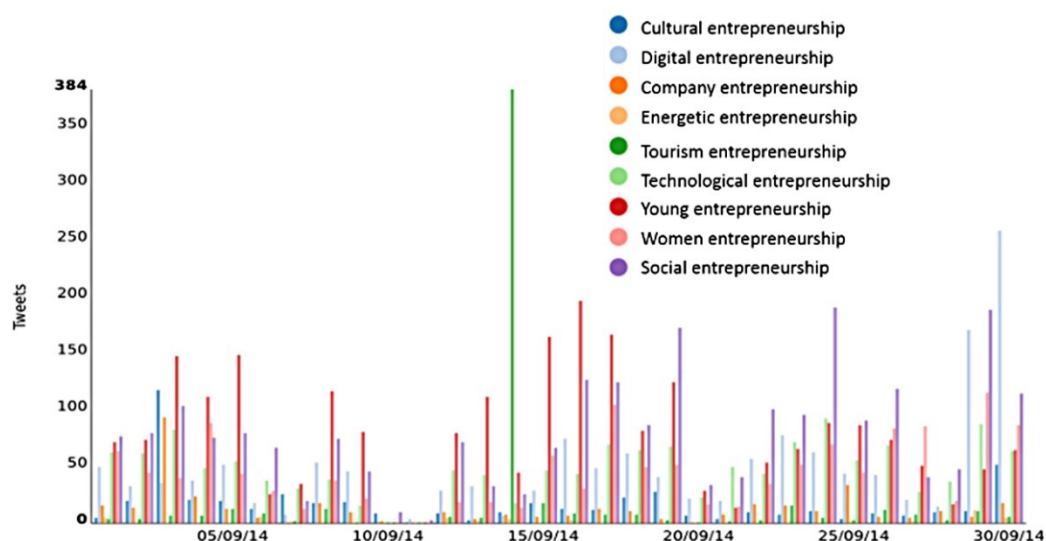


Figura 6:21 Evolución de las menciones de los usuarios sobre productos para el dominio de la industria de la Innovación y el emprendimiento empresarial.

En este sentido, el turismo es la marca de producto que la mayoría de los usuarios han mencionado, con 384,00 tweets, el producto pertenece a la industria de la innovación y el emprendimiento empresarial. La Tabla 6.2 describe los conceptos que apoyan el modelo formal de la Figura 4:12, que utilizan los valores obtenidos para este experimento en el dominio de la industria de la innovación y el emprendimiento empresarial.

Tabla 6:2:

Descripción de los conceptos del modelo formal para el caso de uso aplicado al dominio de la industria de la innovación y el emprendimiento empresarial.

CONCEPTO	VALOR
Usuario	user@Bill_(nombre_ficticio)
Categoría de producto	Industria de la Innovación y emprendimiento
Producto	Turismo
Características de producto	Permanencia / Temporal
Red social	Twitter
Clasificación de red social	Uso global
Características de red social	Microblogging
Categoría	Perfil

La Figura 6:22, muestra la relación de los conceptos definidos en la ontología del sistema que permiten ser aplicados a este caso de estudio, por medio de la cual se muestra el análisis de las características de un producto denominado <<Turismo>>; el producto pertenece a una categoría de tipo <<Empresarial>>. Además, el producto tiene una característica que se llama la <<Innovación>>, en el que además se trata de una característica de interés para el usuario.

En esta prueba, se analizaron 4.624 tweets, de los cuales 571 correspondientes a un setenta y ocho por ciento, mencionan las características de <<Innovación>>, el modelo establece un rasgo de interés por parte del usuario hacia dicho producto y característica, ya que dicha información sería utilizada en futuras recomendaciones a través de un Sistema de Recomendación.

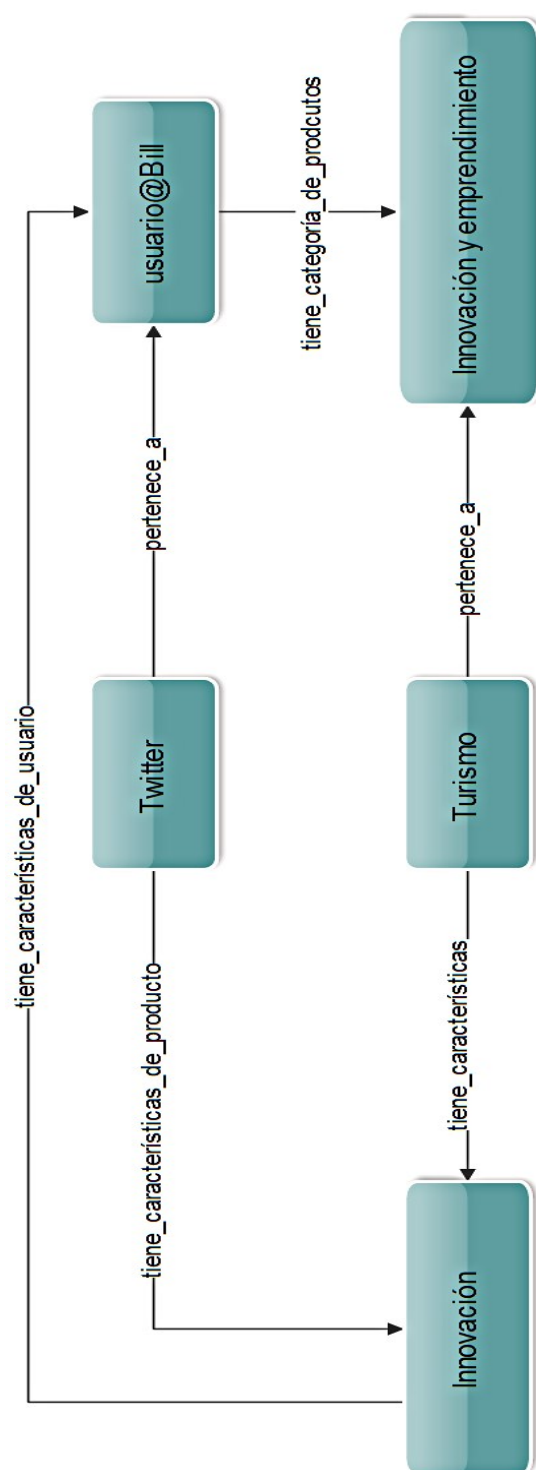


Figura 6:22 Relaciones de características para el dominio de la innovación y el emprendimiento empresarial.

Por lo tanto, la relación del usuario con una red social permite caracterizar la información para un producto. Los resultados del análisis a partir de estos conceptos se muestran en la Figura 6:23, además se apoya con las relaciones establecidas en la Figura 4:12 del modelo formal. Además, sobre la base del modelo de semántica, la información del producto Turismo y sus propias características, tales como la "innovación", se indica. Por lo tanto, esta característica nos permite identificar las preferencias del usuario para la validación de un producto. En este sentido, un usuario habla de la seguridad a través de las menciones en sus tweets, además, el experto expresa que tienen un perfil corporativo, ya que mencionan un producto de la marca Audi. Por lo tanto, en la Figura 6:23, muestra los usuarios con un perfil corporativo que menciona la marca Audi y con características de seguridad.

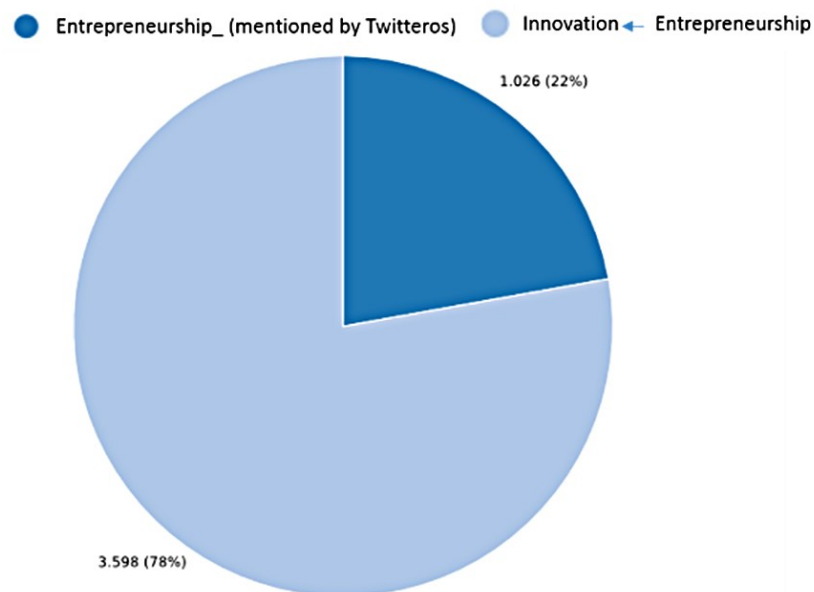


Figura 6:23 Análisis de características mencionadas

En las relaciones de alto nivel, los usuarios de productos y conceptos de redes sociales se detallan. Estos conceptos son la llave que permite que la información obtenida

a partir de datos de la Web pueda ser modelada y validada. Por otra parte, también se representan las relaciones de los conceptos del modelado entre un <<Usuario>> que tiene la red social <<Twitter>> y el nombre del <<Producto>>. Este concepto pertenece a una categoría de tipo de <<Iniciativa digital empresarial>>, el producto también con características específicas, como la <<Innovación y la publicidad>>. Estos conceptos se muestran en la Figura 6:24.

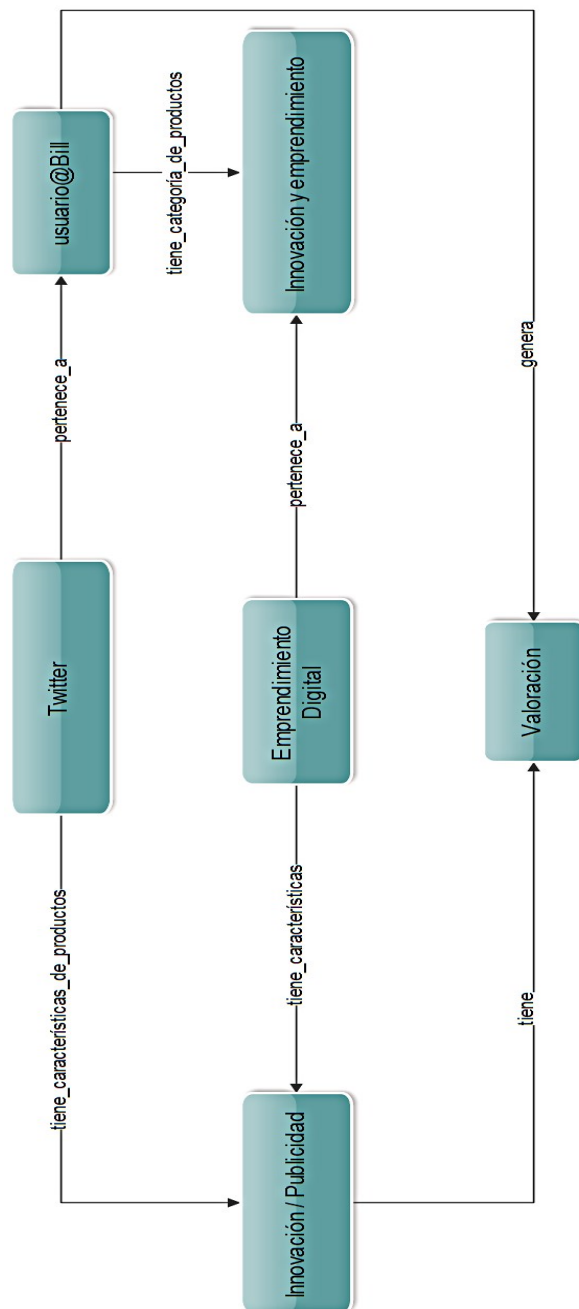


Figura 6:24 Relación de conceptos para generar la validación del caso de estudio

En esta prueba, el experimento permite extraer los tweets de los usuarios que mencionan el producto <<Emprendimiento empresarial digital>>, sus características de <<Innovación>> y <<Publicidad>>. El usuario también tiene una relación con la valoración de estas características. A fin de satisfacer el interés de los usuarios, y de modo que en el futuro la información servirá para hacer recomendaciones a través de un sistema de recomendación, la información se almacena en la ontología propuesta. Varios enfoques pueden ser aplicados a la valoración de dichos análisis y recomendaciones.

Por ejemplo, se puede basar en los resultados de sentimiento que permiten conocer las opiniones de los usuarios en relación con una clasificación de productos a través. Por lo tanto, esta información sirve para caracterizar semánticamente las marcas y productos, así como que sea posible incorporar nuevos conceptos que pueden ser utilizados por la ontología sobre los hashtags de los tweets. La misma figura refleja la relación entre los conceptos de <<Usuario>>, <<Rating>> y los conceptos de <<Innovación / publicidad>> para este experimento.

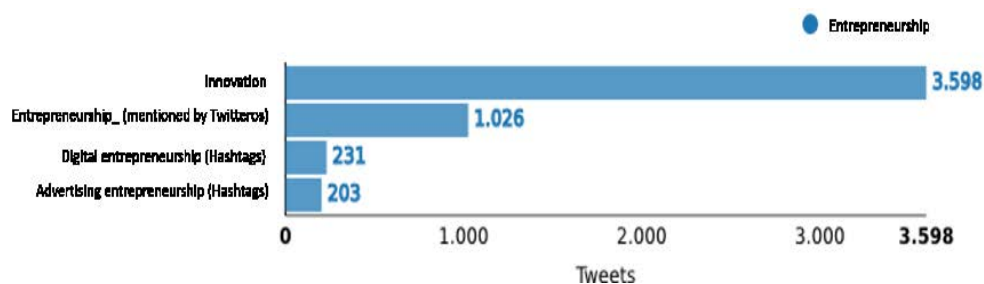


Figura 6:25 Análisis de tweets que mencionan distintas características en el dominio empresarial.

La Figura 6:25, muestra los resultados obtenidos a partir del análisis de los tweets que se mencionan por los usuarios. En esta prueba, se identifican 231 tuits que mencionan una relación con la iniciativa empresarial digital y las características de innovación y la publicidad.

Al igual que en el caso de estudio descrito en la Sección 6.3.1. En este caso de estudio se reiteran la comprobación de las hipótesis planteadas en la presente investigación. Asimismo, se realiza un experimento que permite analizar la ontología del dominio empresarial, sus relaciones y sus conceptos definidos en el modelo ontológico descrito en la Sección 4.

La Figura 6:26, muestra un fragmento de la taxonomía de la ontología para el modelado en el dominio de la industria de la innovación y el emprendimiento empresarial.

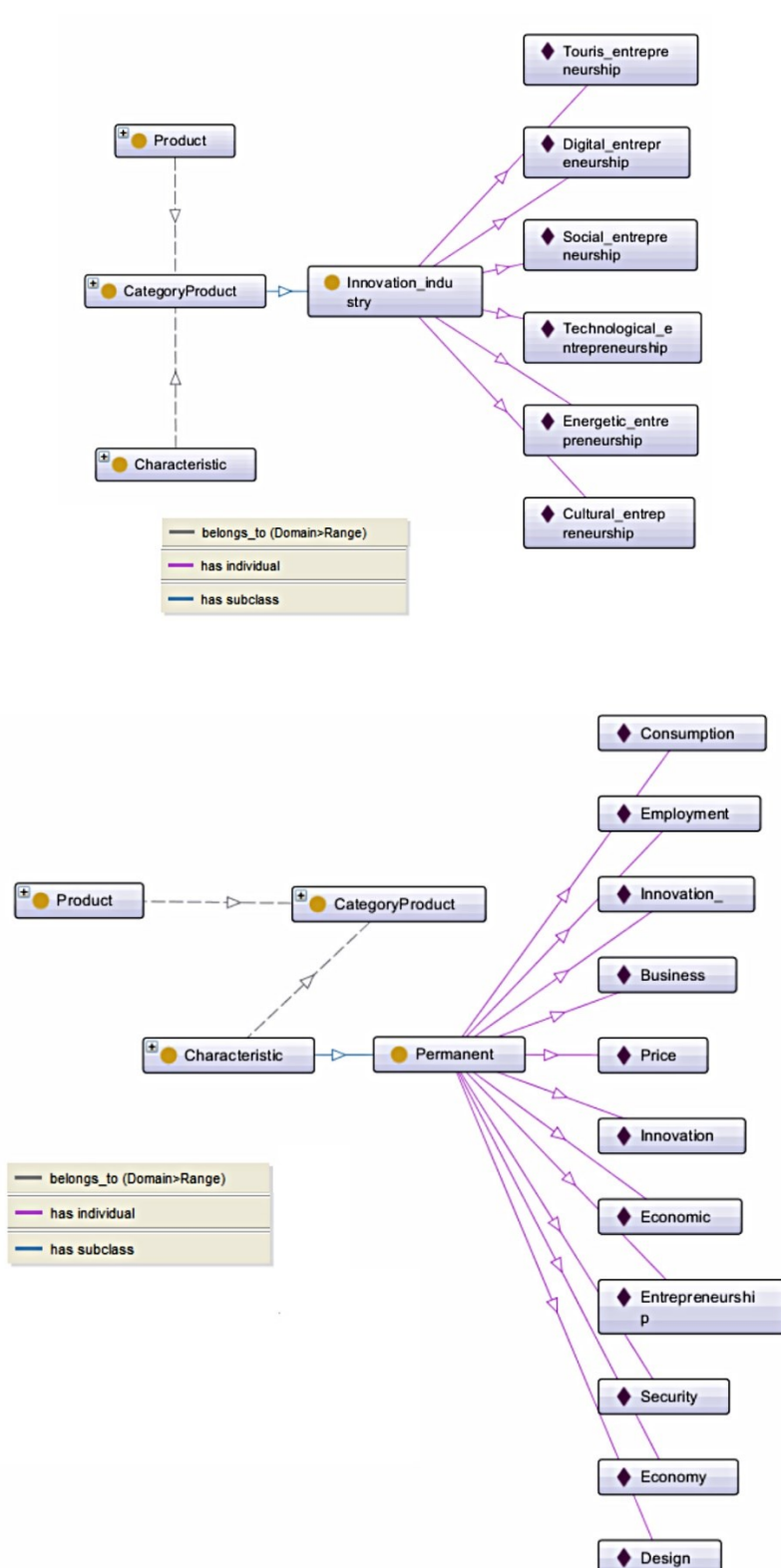


Figura 6:26 Fragmento de la taxonomía en el dominio de la industria de la innovación y el emprendimiento empresarial.

6.4. Validación del modelo utilizando la herramienta de recomendación.

6.4.1. Recomendaciones en el dominio de la industria automotriz

La herramienta de recomendaciones permite validar, por un lado, el diseño de la arquitectura propuesta en el Capítulo 5. Por otro lado, la generación de recomendaciones en distintos dominios. En esta sección, se describe el uso de la herramienta en el dominio de la industria automotriz.

En este sentido, la herramienta de recomendación utiliza el almacenamiento de datos o base de conocimiento propuesta y descrita en el apartado de la Sección 5.3.4. A su vez, la herramienta es utilizada por la arquitectura propuesta, que a su vez, es basada en el modelo conceptual multi-dominio propuesto en esta investigación y descrito en el Capítulo 4. De esta manera, en la validación de las recomendaciones, la herramienta recomienda productos, para este caso, recomienda las marcas de automóviles y sus características que se desean ofrecer hacia los usuarios, además que mantienen una relación con las características de los usuarios. Por lo tanto, el cruce de las características de los usuarios y las características de los productos entrantes al sistema se traducen en otras más técnicas que ponderan que productos son los más acordes a la petición.

Por otra parte, la herramienta de recomendación permite al usuario acceder a un tipo de producto, de forma que el usuario podrá interactuar con el sistema para gestionar los productos, sus características y servicios. En este sentido, la información que se almacena en la ontología es representada en forma de una base de datos. Por lo tanto, la información permite realizar recomendaciones siempre y cuando sean posibles de llevar a cabo. La Figura 6:27 muestra el acceso a un tipo de producto, para este caso, las marcas del dominio de la industria del automóvil.

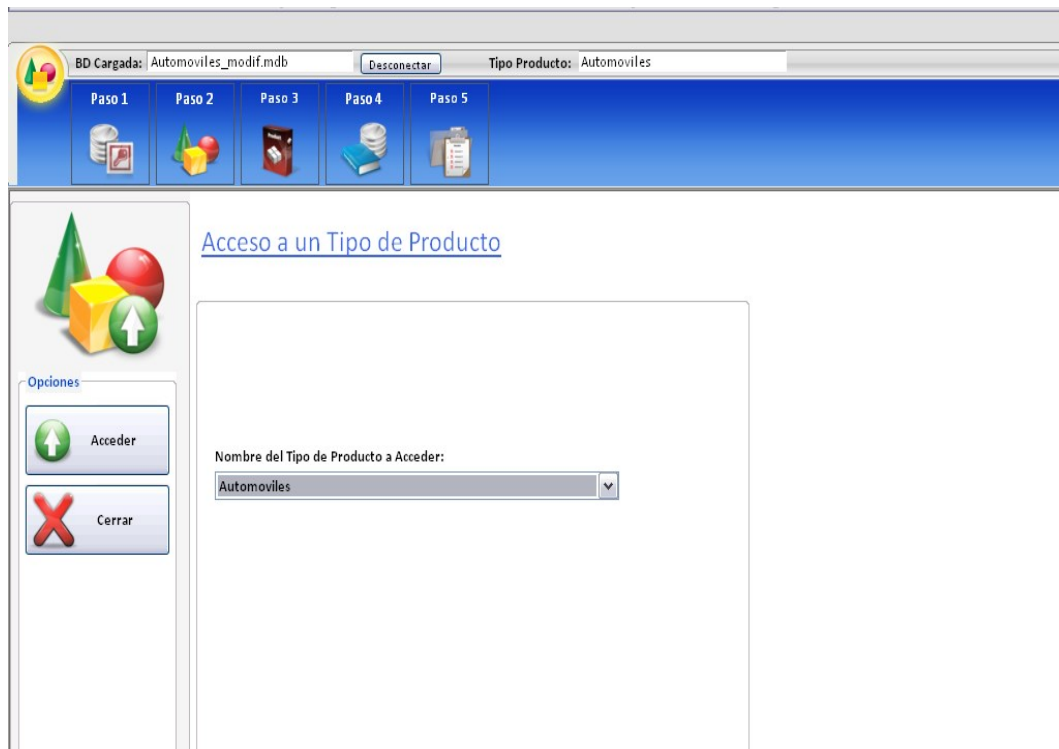


Figura 6:27 Herramienta de recomendación, interfaz de acceso a un producto del dominio de la industria automotriz

El almacenamiento de datos proporciona información detallada sobre las características de productos que serán relevantes para el usuario y por consiguiente, son tratadas como características que mantienen una relación con el usuario.

La herramienta de recomendación proporciona una ventana para gestionar y definir los pesos que ponderan las relaciones entre las características de productos y las características de usuarios. La forma de introducir los pesos es a través de una lista con valor numérico de ponderación y una breve descripción pudiendo introducir tantos como se consideren oportunos. La Figura 6:28, muestra la ventana para la confirmación de pesos sobre los productos en el dominio de la industria automotriz.



Configuración de Pesos

Opciones

 Guardar

 Cerrar

Pesos:

	Valor	Descripción
	-1	Malo
	0	Regular
	1	Normal
	2	Bueno
	3	Excelente
		

Figura 6:28 Ponderación en el dominio de la industria automotriz

En este sentido, existe la necesidad de poder gestionar las relaciones que existen entre las características de producto y características de usuario para un tipo de producto determinado. Para este caso, un producto determinado Ford, con características de seguridad, información que se encuentra almacenada en la Base de datos. Por lo tanto, la herramienta permite seleccionar dentro de una maya las características relacionables de usuario-producto que se desee unir con un peso. Esta operación se podrá repetir tantas veces como relaciones se desee definir.

La Figura 6:29 muestra la maya con las relaciones para las características del producto en el dominio de la industria automotriz y las características de un usuario.

BD Cargada: Automoviles_modif.mdb Desconectar Tipo Producto: Automoviles

Paso 1 Paso 2 Paso 3 Paso 4 Paso 5

Relación Características Producto-Usuario

Opciones: Guardar Cerrar

		Categoría NCAP			
		1	2	3	4
T_DISEÑO	(L)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	(LP)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	(T)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	X	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
T_OPCION	ECO	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	ESP	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	GPS	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	SPORT	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	TV	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
T_PRECIO	(Alta)	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	(Baja)	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
	(Indiferente)	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	(Media)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Figura 6:29 Relaciones entre características de usuarios y características de productos en el dominio de la industria automotriz

Un listado de relaciones sobre los productos, por un lado, permite gestionar las relaciones existentes en el sistema para un determinado producto. Por otro lado, representan las características relacionables del sistema. La herramienta de recomendación permite definir recomendaciones a partir de un formulario de petición que presenta la lista de características que ofrece el sistema. Una vez rellenado el formulario, se valida su contenido y solicita los productos recomendados.

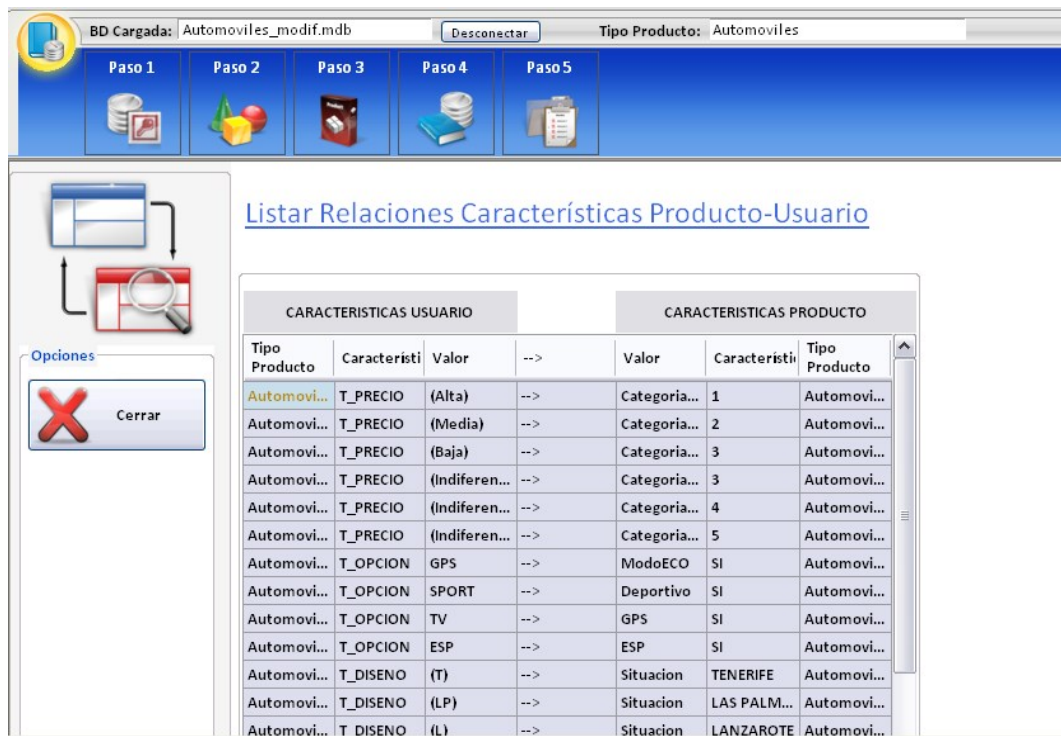


Figura 6:30 Lista de características de productos

El resultado que se presenta en la pantalla de la Figura 6:30, es el que resulta del formulario petición de recomendación descrito en el punto anterior. Los productos que se visualizan son aquellos que el sistema prevé más acordes a las necesidades/gustos descritos en la petición y según los parámetros inscritos en el sistema en el momento de realizar la petición. La herramienta además proporciona la función de exportación de recomendaciones en formato HTML y formato XLS.

Finalmente, la Figura 6:31 muestra la ventana para realizar la petición de la recomendación. En consecuencia, la Figura 6.32 muestra la ventana con la lista de las características seleccionadas en la petición de la recomendación. Por último, la Figura 6.33 muestra la respuesta sobre los productos relacionados con el dominio de la industria automotriz recomendados hacia el usuario en un formato Web.

BD Cargada: Automoviles_modif.mdb Desconectar Tipo Producto: Automoviles

Paso 1 Paso 2 Paso 3 Paso 4 Paso 5

Realizar Petición de Recomendación

Opciones

Aceptar

ver HTML

Cerrar

Características

T_DISEÑO (Radio)

☒ (L)

☐ (LP)

☐ (T)

☐ X

T_OPCION (Check)

☐ ECO

☐ ESP

☐ GPS

☐ SPORT

☐ TV

T_PRECIO (Radio)

☒ (Alta)

☐ (Baja)

☐ (Indiferente)

Figura 6:31 Petición de recomendación

BD Cargada: Automoviles_modif.mdb Desconectar Tipo Producto: Automoviles

Paso 1 Paso 2 Paso 3 Paso 4 Paso 5

Recomendaciones

Opciones

Vision Web

Visión Excel

Cerrar

Petición

	Característica	Valor	Característica	Valor
▶	T_DISEÑO	(T)		
	T_PRECIO	(Alta)		
	T_SEGURIDAD	()		
	T_DISEÑO	(T)		

Recomendaciones

	Nombre	CategoríaNCAP	Foto	ModoECO	Libres
▶	Opel Insignia 1...	3	.\Fotos\Opelln...	SI	10
	Honda Civic 12...	3	.\Fotos\Honda...	SI	5
	Mercedes Clas...	4	.\Fotos\Merce...	SI	20
	Audi A3 140CV	4	.\Fotos\AudiA...	SI	7
	Seat Leon FR 1...	5	.\Fotos\LeonF...	SI	1

Figura 6:32 Lista de características de productos

Solicitud y Resultado de Recomendaciones para Automoviles

AMBITO: Automoviles RECOMENDACIONES: 5

Datos Peticion

CARACTERÍSTICA	VALOR
T_DISEÑO	(T)
T_PRECIO	(Alta)
T_SEGURIDAD	()

Productos Recomendados

NOMBRE	CATEGORIA	NCAP	FOTO	MODOE	CO	LIBRES	DEPORTIVO	GPS	SITUACION	ESP	COD	DESCRIPCION
Opel Insignia 140CV Excellence	3		./Fotos/OpelInsignia.jpg	SI		10	SI	SI	TENERIFE	SI	1	Nuevo Opel Insignia Edicion 2014,
Honda Civic 120CV Soort	3		./Fotos/HondaCivic.jpg	SI		5	SI	SI	TENERIFE	SI	2	Honda Civic 2015 acabado Soort

Figura 6:33 Resultado de recomendaciones en el dominio de la industria automotriz

6.4.2. Recomendaciones en el dominio de la industria de la innovación y el emprendimiento empresarial

Con el fin de remarcar la validación del modelo propuesto utilizando una herramienta de recomendación integrada en la arquitectura propuesta. A continuación, se realiza la validación del modelo para un dominio distinto. En este caso, se presenta una validación en el dominio de la industria de la innovación y el emprendimiento empresarial. Ahora bien, los conceptos de la herramienta que han sido descritos con mayor detalle anteriormente en la Sección 6.3.2, se mantienen de forma permanente. Es decir, que solo se cambian los productos y características para el dominio a recomendar.

Por lo tanto, para este dominio se utiliza la información almacenada en la base de conocimientos o almacenamiento de datos. En la Figura 6:34 se muestra el acceso a un tipo de producto, para este caso, los productos del dominio de la industria de la innovación y el emprendimiento empresarial.

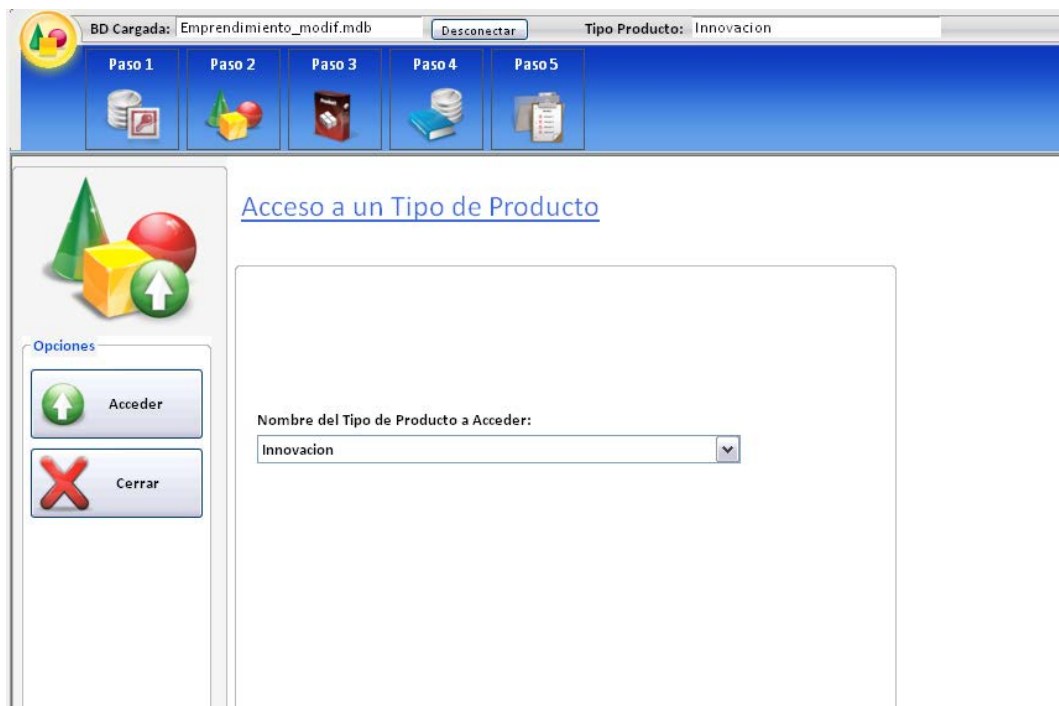


Figura 6:34 Herramienta de recomendación, interfaz de acceso a un producto del dominio de la industria de la innovación y el emprendimiento empresarial

Por consiguiente, se gestionan las relaciones que existen entre las características de producto y las características de usuario para un tipo de producto determinado. Para este caso, un producto determinado “turismo”, dicho producto mantiene características de innovación, información que se encuentra almacenada en la base de datos. Por lo tanto, de forma similar al caso del dominio automotriz presentado con anterioridad, la herramienta permite seleccionar dentro de una maya las características relacionables de usuario-producto que se desee unir con un peso. Asimismo, esta operación se podrá repetir tantas veces como relaciones se desee definir. La Figura 6:35 muestra la maya con las relaciones para las características del producto en el dominio de la industria de la innovación y el emprendimiento empresarial y las características de un usuario. Finalmente, la herramienta genera las recomendaciones determinadas, la Figura 6:36 muestra la ventana que permite visualizar las recomendaciones de productos en el dominio de la industria de la innovación y el emprendimiento empresarial.

BD Cargada: Emprendimiento_modif_dany Desconectar Tipo Producto: Innovacion

Paso 1 Paso 2 Paso 3 Paso 4 Paso 5

Relación Características Producto-Usuario

Opciones

Guardar

Cerrar

		1	2	3	4
T_CULTURAL	ECO	<input checked="" type="checkbox"/> NORMAL	<input checked="" type="checkbox"/> BUENO	<input checked="" type="checkbox"/> BUENO	<input checked="" type="checkbox"/> BUENO
	ESP	<input checked="" type="checkbox"/> NO MUY BUENO	<input checked="" type="checkbox"/> MUY BUENO	<input checked="" type="checkbox"/> MUY BUENO	<input checked="" type="checkbox"/> MUY BUENO
	GPS	<input checked="" type="checkbox"/> BUENO	<input checked="" type="checkbox"/> NO MUY BUENO	<input checked="" type="checkbox"/> BUENO	<input checked="" type="checkbox"/> MUY BUENO
	SPORT	<input checked="" type="checkbox"/> INDIFERENTE	<input checked="" type="checkbox"/> BUENO	<input checked="" type="checkbox"/> MUY BUENO	<input checked="" type="checkbox"/> MUY BUENO
	TV	<input checked="" type="checkbox"/> INDIFERENTE	<input checked="" type="checkbox"/> NORMAL	<input checked="" type="checkbox"/> INDIFERENTE	<input checked="" type="checkbox"/> BUENO
T_DIGITAL	(L)	<input checked="" type="checkbox"/> MUY BUENO	<input checked="" type="checkbox"/> MUY BUENO	<input checked="" type="checkbox"/> NORMAL	<input checked="" type="checkbox"/> BUENO
	(LP)	<input checked="" type="checkbox"/> MUY BUENO	<input checked="" type="checkbox"/> MUY BUENO	<input checked="" type="checkbox"/> BUENO	<input checked="" type="checkbox"/> BUENO
	(T)	<input checked="" type="checkbox"/> BUENO	<input checked="" type="checkbox"/> NORMAL	<input checked="" type="checkbox"/> INDIFERENTE	<input checked="" type="checkbox"/> NORMAL
	X	<input checked="" type="checkbox"/> MUY BUENO	<input checked="" type="checkbox"/> NORMAL	<input checked="" type="checkbox"/> INDIFERENTE	<input checked="" type="checkbox"/> INDIFERENTE
T_TECNOLOGICO	()	<input checked="" type="checkbox"/> MUY BUENO	<input checked="" type="checkbox"/> INDIFERENTE	<input checked="" type="checkbox"/> BUENO	<input checked="" type="checkbox"/> BUENO
	(i)	<input checked="" type="checkbox"/> BUENO	<input checked="" type="checkbox"/> NORMAL	<input checked="" type="checkbox"/> MUY BUENO	<input checked="" type="checkbox"/> BUENO
	(O)	<input checked="" type="checkbox"/> MUY BUENO	<input checked="" type="checkbox"/> MUY BUENO	<input checked="" type="checkbox"/> BUENO	<input checked="" type="checkbox"/> MUY BUENO
	(P)	<input checked="" type="checkbox"/> BUENO	<input checked="" type="checkbox"/> BUENO	<input checked="" type="checkbox"/> INDIFERENTE	<input checked="" type="checkbox"/> NORMAL
	(V)	<input checked="" type="checkbox"/> NORMAL	<input checked="" type="checkbox"/> MUY BUENO	<input checked="" type="checkbox"/> NORMAL	<input checked="" type="checkbox"/> MUY BUENO

Figura 6:35 Relaciones entre características de usuarios y características de productos en el dominio de la industria de la innovación y el emprendimiento empresarial

Solicitud y Resultado de Recomendaciones para Innovacion

AMBITO Innovacion RECOMENDACIONES 5

Datos Peticion

CARACTERÍSTICA	VALOR
T_CULTURAL	ESP
T_DIGITAL	(L)
T_TECNOLOGICO	()
T_TURISMO	(Alta)

Productos Recomendados

NOMBRE	CATEGORIA	FOTO	TECNOLOGICO	LIBRES	DIGITAL	ENERGIA	TURISTICO	SOCIAL	COD	DESCRIPCION
Turistico	5	.\Fotos\Innova_Turistico.jpg	NO	1	SI	NO	ESPANA	SI	8	Emprendimiento Turistico 2014

Figura 6:36 Resultado de recomendaciones en el dominio de la industria de la innovación y el emprendimiento empresarial

6.5. Conclusión del caso de estudio y de la aplicación del modelo

El modelo conceptual propuesto en esta tesis es aplicado a distintos dominios a partir de dos casos de estudio planteados en este capítulo. En el proceso de validación y a través de estas pruebas, por un lado, el modelo permite la representación del conocimiento desde distintos dominios a partir de la información extraída de los medios sociales. Por otro lado, permite la caracterización de un contenido estructurado sobre los distintos productos, marcas, sus características y servicios ofertados en las redes sociales, a su vez, puede modelar y gestionar el conocimiento de diferentes perfiles de usuarios, productos y medios sociales caracterizados en múltiples dominios. En consecuencia, esto comprueba cada una de las hipótesis planteadas al inicio de la investigación, ya que a partir del diseño de un modelo y su aplicación para la representación del conocimiento en dos casos de estudio, se comprueban cada una de las hipótesis.

Los conceptos definidos para la aplicación del modelo en este capítulo, por un lado, permiten la representación del conocimiento para diversos dominios a partir de la información extraída de los medios sociales, para este caso, de la red social Twitter. Por otro lado, la definición de una ontología permite la caracterización de un contenido sobre los distintos productos que satisfacen las necesidades de los usuarios.

Por otra parte, la definición y descripción de cada uno de los componentes que integran el diseño del marco computacional de trabajo propuesto y descrito en el Capítulo 5. Permite aplicar el modelo conceptual propuesto comprobando la capacidad que brindan tanto el modelo, como los componentes propuestos para un soporte en el funcionamiento del sistema.

6.6. Sumario

En este capítulo se han presentado dos casos de estudio que permiten por un lado, validar el modelo conceptual propuesto en la presente tesis doctoral. Por otro lado, validar las cuatro hipótesis planteadas al inicio de la investigación.

La primera y segunda hipótesis planteadas suponían el diseño de un modelo conceptual multi-dominio para recomendaciones, que sirva como una base para la operabilidad de Sistemas Basado en Conocimiento. Además, el modelo propuesto podría representar la caracterización de productos, medios sociales y perfiles de usuarios. Asimismo, gestionar el conocimiento. Por otra parte, a partir del modelo conceptual propuesto la información es extraída de los medios sociales, a su vez, la creación de una ontología que define el modelo se representa en forma de una base de datos relacional que es utilizada por la herramienta de recomendación para la generación de recomendaciones de productos en distintos dominios. Estas hipótesis se han comprobado a partir de la definición de los conceptos propuestos en el Capítulo 4 y el diseño de una arquitectura para la solución propuesta en el Capítulo 5.

La tercera y cuarta hipótesis planteaban el diseño de una plataforma que integrara el modelo conceptual, además de la creación de un modelo conceptual genérico, adaptable a cualquier dominio de contenido estructurado y conocimiento semántico. Estas hipótesis se han validado a partir del diseño de un marco de trabajo que permite aplicar el modelo para distintos dominios. De forma que, en la aplicación del modelo en dos casos de estudio propuestos en este capítulo permiten comprobar las hipótesis que enmarcan la investigación. La información extraída de los medios sociales, para estos casos, de la red social Twitter genera resultados relevantes en la información de los productos y características que los usuarios mencionan. En consecuencia, la definición de la ontología permite la caracterización de un contenido sobre los distintos productos que satisfacen las necesidades de los usuarios.

Los dos casos de estudio planteados en este capítulo, por un lado permiten representar un conocimiento a partir de la extracción de la información semántica en los medios sociales, por otro lado, la caracterización de distintos productos, sus

características y servicios ofertados en las redes sociales, el proceso se realiza a través de la red social Twitter que permite filtrar la información para la caracterización de los contenidos infiriendo el conocimiento previo a un contexto de información en cada caso de estudio. Además, el análisis realizado en cada caso de estudio sobre las menciones que realizan los usuarios hacia los productos, permite modelar y gestionar los contenidos. Con estas hipótesis se ha comprobado la capacidad del modelo conceptual propuesto en esta investigación, ya que el modelo puede ser aplicado a distintos dominios de contenido estructurado y semántico, asimismo permite representar el conocimiento desde diferentes dominios en los medios sociales.

Capítulo 7

Conclusiones y futuras líneas de investigación

En este capítulo se presentan las conclusiones finales de la investigación que se han expuesto en la presente tesis doctoral. Junto con el repaso de las principales aportaciones que incluye la investigación de esta tesis. Asimismo, se propone un conjunto de futuras líneas de investigación que pueden ser una continuación de la misma. Finalmente, se lista el conjunto de publicaciones científicas que se han realizado a partir de la presente investigación.

7.1. Conclusiones

Al inicio de la presente tesis doctoral se describió la problemática que se abordaría en su momento de la investigación.

El modelo propuesto permite la representación del conocimiento mediante técnicas de recuperación de información en los medios sociales. Los conceptos que se describen en el modelo pueden generar una caracterización de productos y perfiles de usuario basado en el modelado de contenido a partir de una red social. Además, el modelo conceptual es válido para diferentes dominios, tales como, la industria automotriz y la industria del emprendimiento.

El modelo permite ser transferido a un dominio diferente para modelar diversos productos, lo que permite la extracción de contenido y el establecimiento de una base de conocimientos. Asimismo, establece la generación de recomendaciones para los usuarios a partir del contenido caracterizado respecto a cualquier producto o servicio, partiendo de un SR. Por otro lado, el modelo propuesto es basado en una ontología que define el modelo conceptual multi-dominio en forma de una base de datos relacional descrita en el Capítulo 4. Dicha ontología permite la gestión de la información extraída de los medios sociales y, a su vez, establece las relaciones entre los conceptos para una caracterización de los perfiles de usuario, los productos y las redes sociales. Asimismo, el modelo es capaz de representar el contenido de un producto extraído de red social preferida del usuario. Para los estudios de los casos de estudio planteados en esta investigación, el experto del dominio ha seleccionado un conjunto de productos para determinar las características finales sobre la base de las referencias del usuario.

Además, este modelo permite la anotación semántica de palabras claves sobre las entidades o conceptos que conforman la ontología. Esta información es utilizada como preferencias para futuras recomendaciones. En la validación realizada, los autores han comprobado, por un lado, que el modelo es compatible con la representación de los elementos necesarios para recomendaciones de los distintos productos ofertados en la Web y sus características con los usuarios potenciales. De esta forma, se generarán

recomendaciones basadas en la recuperación de información a partir de la información extraída desde las redes sociales. Por otro lado, la comprobación de las hipótesis planteadas en la investigación.

En consecuencia, se podrían hacer recomendaciones dinámicas basadas en la continua evolución del usuario y la información sobre los productos publicados en los medios de comunicación social.

En los distintos casos de estudio presentados en este capítulo, se reafirma que la información se ha extraído de una red social de confianza Twitter. Además, el análisis sobre las estadísticas extraídas y presentadas en esta investigación apoya los conceptos y relaciones definidas en el modelo ontológico propuesto en esta tesis doctoral. De manera que, el modelo propuesto es capaz de representar el conocimiento de distintos dominios. Asimismo, el modelo conceptual puede caracterizar un contenido estructurado sobre los distintos productos, marcas, sus características y servicios ofertados en las redes sociales, a su vez, puede modelar y gestionar el conocimiento de diferentes perfiles de usuarios, productos y medios sociales caracterizados en múltiples dominios. Por lo tanto, la información almacenada en la base de conocimiento sirve para la generación de futuras recomendaciones. Actualmente, existen recursos que permiten conocer la opinión del usuario a ser conocido con respecto a las calificaciones de algunos productos a través de la positividad y la negatividad de las palabras utilizadas en un dominio. Pero en esta investigación se ha optado por la positividad y objetividad de las menciones que el usuario realiza a través de sus tweets. De manera que el sistema utiliza reglas que caracterizan positiva y objetivamente las menciones. En el diseño de la arquitectura es posible realizar un análisis de sentimiento a partir de las menciones que los usuarios realizan sobre los productos y sus características.

En el proceso de recomendación, los componentes que integran el diseño del marco computacional de trabajo hacen del proceso una solución al problema, ya que a partir de la utilización de una herramienta de recomendación, permite generar las recomendaciones de productos basadas en el contenido del modelo conceptual propuesto para esta tesis doctoral.

7.2. Futuras líneas de investigación

El modelo conceptual multi-dominio permite la representación del conocimiento mediante el filtrado de información semántica en los medios sociales. Por lo tanto, a partir de la representación del conocimiento es posible añadir nuevos conceptos al modelo conceptual multi-dominio que permitan mejorar la representación del conocimiento desde distintos paradigmas.

La información almacenada en la Web permite trasladar el modelo conceptual propuesto en esta investigación, hacia una aplicación distinta a los dominios que se han presentado en esta tesis doctoral, lo que permita la interoperabilidad del modelo entre otros dominios distintos a los presentados en esta investigación.

Asimismo, el modelo conceptual multi-dominio utiliza la red social Twitter para la extracción de la información, como línea futura, se aplicaría el modelo para la extracción semántica y de contenido estructurado a una red social distinta a Twitter. Por ejemplo, la red social Facebook, que mediante la utilización de las Tecnologías Semánticas serviría como una nueva fuente de información para la extracción de contenido.

7.3. Publicaciones realizadas a raíz de esta tesis doctoral

Como consecuencia de esta investigación y, a la formación que se pretende adquirir al programa de doctorado, se han realizado las siguientes publicaciones en diferentes ámbitos científicos:

- Publicaciones en revistas con factor de impacto.
 - Daniel Villanueva-Vasquez, Israel González-Carrasco, José L. López-Cuadrado and Nora Lado. SMORE: Towards a Semantic Modeling for knowledge representation on social media. *Science of Computer Programming*. Accepted Original Article (June 2015).
 - Posada-Gómez, R., Villanueva, D., González, I., García, Á. Aguilar-Lasserre, A. A. and Martínez-Sibaja, A. (2014), Toward an Automatic Parameterization System for the Classification of Persian Lemons Using Image-Processing Techniques. *Journal of Food Process Engineering*. Doi: 10.1111/jfpe.12164. Original Article.
 - Villanueva, D., Posada, R., González, I., García, Á. and Martínez, A. (2014), Monitoring of a Sugar Crystallization Process with Fuzzy Logic and Digital Image Processing. *Journal of Food Process Engineering*. Doi: 10.1111/jfpe.12122. Original Article.
- Publicaciones en congresos nacionales e internacionales.
 - Implementación de un operador por Hardware para la estimación del flujo óptico por el método de correlación. JCRA 2014. Jornada de Computación Reconfigurable y Aplicaciones. Valladolid Septiembre 2014 España. Conference Paper.
 - García-Crespo, Á.; Villanueva, D.; González-Carrasco, I.; López-Cuadrado, J.L. Chamorro, J.M.; Iglesias, I.; Celorio, C. EPSIS: Entretenimiento y Publicidad Segmentada en Entornos Inmersos. CTVDI Congreso Internacional de Televisión Digital Interactiva, CTVDI-13, 29-31 Octubre 2013. Mallorca, España. Conference Paper.

- Lemos, M. L., Vasquez, D. V., Radzimski, M., Lemos, A. L., & Gómez-Berbís, J. M. (2012). RING: A Context Ontology for Communication Channel Rule-based Recommender System. *SeRSy 2012*, 73. Workshop Paper.
- Otros trabajos realizados durante el periodo de investigación.
 - Daniel Villanueva-Vasquez, Israel González-Carrasco, Ángel García-Crespo, José L. López-Cuadrado. “*A Review of Social Media Networks on Recommender Systems Models*”. *IETE Technical Review* Original Article (Under Review)
 - Daniel Villanueva-Vasquez, Israel González-Carrasco, Ángel García-Crespo, José I. López-Cuadrado “*SocialMent: A Framework for Support to Analysis and Management Big Data on Social Media*”. *Information Retrieval Journal* Original Article (Under Review)
 - Daniel Villanueva-Vasquez, Israel González-Carrasco, Ángel García-Crespo, José L. López-Cuadrado. “*RESKISOM: Representation of semantic knowledge in Social Media*”. An approach using information retrieval by recommendations. *Journal of Universal Computer Science* Original Article (Under Review)
 - Daniel Villanueva, Israel González, Ángel García and José I. López. *FRAPER: Towards a multi-domain Framework for Personalized Recommendations in Social Media*. *ESWA* Original Article (Under Review)

- Ángel García, Israel González, José L. López, Daniel Villanueva and Álvaro Muñoz. *Empowering Smart Cities with a framework for creating Cultural Entertainment Systems for the IoT. Transactions on Multimedia Computing Communications and Applications* Original Article (Under Review)

Bibliografía

- Abel, F, C Hauff, and GJ Houben (2012). "Semantics+ filtering+ search= twitcident. exploring information in social web streams." In *Proceedings of the 23rd ACM conference on Hypertext and social media* (pp. 285-294). ACM. <http://dl.acm.org/citation.cfm?id=2310043>
- Adomavicius, G., and A. Tuzhilin (2005). "Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions." *IEEE Transactions on Knowledge and Data Engineering* 17(6): 734-749. <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1423975>
- Advani, A, SW Tu, and MA Musen (1997). "Domain modeling with integrated ontologies: Principles for reconciliation and reuse." *AMIA ANNUAL FALL SYMPOSIUM*. http://www.researchgate.net/publication/25901384_Domain_Modeling_with_Integrated_Ontologies_Principles_for_Reconciliation_and_Reuse/file/79e4150c74de780f31.pdf
- Agichtein, E, C Castillo, and D Donato (2008). "Finding high-quality content in social media." *Proceedings of the 2008* <http://dl.acm.org/citation.cfm?id=1341557>
- Agius, HW, and MC Angelides (1999). "Developing knowledge-based intelligent multimedia tutoring systems using semantic content-based modelling." *Artificial Intelligence Review* (Kanade 1996): 55-83. <http://link.springer.com/article/10.1023/A:1006569626086>
- Agogino, AM, S Srinivas, and KM Schneider (1988). "Multiple sensor expert system for diagnostic reasoning, monitoring and control of mechanical systems." *Mechanical Systems and Signal Processing*, 2(2), 165-185. <http://www.sciencedirect.com/science/article/pii/0888327088900416>
- Alba, J. (2007). "La Web Semántica." *Bit*. <http://dialnet.unirioja.es/servlet/articulo?codigo=2334033&orden=126354&info=link>
- Allan, J. (2004). "NLP for IR-Natural Language Processing for Information Retrieval." *Proceedings of the Twelfth Text Retrieval Conference*. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.34.8740>

-
- Alonso, JM, A Munoz, and JA Botia (2008). "Uso de ontologías para facilitar las tareas de extracción y representación de conocimiento en el diseño de sistemas basados en reglas borrosas." In *XIV Spanish ESTYLF conference on fuzzy logic and technologies* (pp. 233-240). <http://ants.inf.um.es/~amunoz/docs/AlonsoMBMS-Estylf08.pdf>
- Andersson, Birger, Maria Bergholtz, Ananda Edirisuriya, Tharaka Ilayperuma, Eric Dubois, Sven Abels, Axel Hahn, Benkt Wangler, and Hans Weigand. (2006). "Towards a Reference Ontology for Business Models *." *Springer*. 482–496
- Ardissono, L, A Goy, R Meo, and G Petrone (1999). "A configurable system for the construction of adaptive virtual stores." *World Wide Web*. <http://link.springer.com/article/10.1023/A:1019252823387>
- Arpírez, JC, and A Gómez-Pérez (2000). "Reference ontology and (ONTO) 2 agent: the ontology yellow pages." *Knowledge and Information Systems*, 2(4), 387-412. <http://link.springer.com/article/10.1007/PL00011649>
- Astrova, I, N Korda, and A Kalja (2007). "Storing OWL ontologies in SQL relational databases." *International Journal of Electrical, Computer and Systems Engineering*, 1(4), 242-247. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.115.6731&rep=rep1&type=pdf>
- Asur, S, and BA Huberman (2010). "Predicting the future with social media." In *Web Intelligence and Intelligent Agent Technology (WI-IAT), 2010 IEEE/WIC/ACM International Conference on* (Vol. 1, pp. 492-499). IEEE. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5616710
- Aurnhammer, M, P Hanappe, and L Steels (2006). "Integrating collaborative tagging and emergent semantics for image retrieval." In *Proc. of the Collaborative Web Tagging Workshop (WWW'06)*. <http://www.ra.ethz.ch/cdstore/www2006/www.rawsugar.com/www2006/17.pdf>
- Baader, F. (2003). "The description logic handbook: theory, implementation, and applications." Cambridge university press. <https://books.google.es/books?hl=es&lr=&id=riSeOKw5l6sC&oi=fnd&pg=PR14&dq=The+Description+Logics+Handbook:+Theory,+Implementation+and+Applications%22&ots=T8UzS3Tyf g&sig=kasM6uWyAMh9FGOUlwu-wisoyxs>
- Baeza-Yates, R, and J Ruiz-del-Solar (2004). "Content-based image retrieval and characterization on specific web collections." In *Image and Video Retrieval* (pp. 189-198). Springer Berlin Heidelberg. http://link.springer.com/chapter/10.1007/978-3-540-27814-6_25
- Balaban, Mira. (1995). "The F-logic approach for description languages." *Annals of Mathematics and Artificial Intelligence* 15(1): 19–60
-

-
- Balabanović, M, and Y Shoham (1997). "Fab: content-based, collaborative recommendation." *Communications of the ACM*. <http://dl.acm.org/citation.cfm?id=245124>
- Bao, Xinlong, Lawrence Bergman, and Rich Thompson (2009). "Stacking recommendation engines with additional meta-features." In *Proceedings of the third ACM conference on Recommender systems - RecSys '09*, New York, New York, USA: ACM Press, p. 109. <http://dl.acm.org/citation.cfm?id=1639714.1639734>
- Barragáns-Martínez, Ana Belén, Enrique Costa-Montenegro, Juan C. Burguillo, Marta Rey-López, Fernando a. Mikic-Fonte, and Ana Peleteiro (2010). "A hybrid content-based and item-based collaborative filtering approach to recommend TV programs enhanced with singular value decomposition." *Information Sciences* 180(22): 4290–4311. <http://linkinghub.elsevier.com/retrieve/pii/S0020025510003427>
- Bates, D, and M Cohen (2001). "Reducing the frequency of errors in medicine using information technology." *Journal of the American Medical Informatics Association*, 8(4), 299-308. http://scholar.google.es/scholar?q=+Reducing+the+frequency+of+errors+in+medicine+using+information+technology.+The+Journal+of+the+American+Informatics+Medical+Association+&btnG=&hl=es&as_sdt=0,5#0
- Baumann, M. (2009). "Web 3.0: The next step for the Internet." *Information Today*. <http://www.questia.com/magazine/1P3-1730564921/web-3-0-the-next-step-for-the-internet>.
- Bedi, P, H Kaur, and S Marwaha (2007). "Trust Based Recommender System for Semantic Web." *IJCAI*. <http://www.aaai.org/Papers/IJCAI/2007/IJCAI07-430.pdf>
- Beel, J, and S Langer (2013). "The impact of demographics (age and gender) and other user-characteristics on evaluating recommender systems." In *Research and Advanced Technology for Digital Libraries* (pp. 396-400). Springer Berlin Heidelberg. http://link.springer.com/chapter/10.1007/978-3-642-40501-3_45
- Belk, M, P. Germanakos, and N Tsianos (2010). "Adapting Generic Web Structures with Semantic Web Technologies: A Cognitive Approach." In *Proceedings of the 4th International Workshop on Personalized Access, Profile Management, and Context Awareness in Databases, in conjunction with VLDB* (pp. 35-40).. http://www.comp.nus.edu/~vldb2010/proceedings/files/vldb_2010_workshop/PersDB_2010/resources/PersDB2010_6.pdf
- Berners-Lee, T, J Hendler, and O Lassila (2001). "The semantic web." *Scientific american* 284(5): 28–37
-

-
- Berners-Lee, T, and L. Kagal (2008). "The fractal nature of the semantic web." *AI magazine*.
<http://www.aaai.org/ojs/index.php/aimagazine/article/viewArticle/2161>
- Bezerra, Byron L.D., and Francisco de A.T. de Carvalho (2004). "A symbolic approach for content-based information filtering." *Information Processing Letters* 92(1): 45–52.
<http://www.sciencedirect.com/science/article/pii/S002001900400184X>
- Blanco-Fernández, Yolanda, Martín López-Nores, José J. Pazos-Arias, and Jorge García-Duque (2011). "An improvement for semantics-based recommender systems grounded on attaching temporal information to ontologies and user profiles." *Engineering Applications of Artificial Intelligence* 24(8): 1385–1397. <http://dx.doi.org/10.1016/j.engappai.2011.02.020>
- Boardman, G, and H Lu (2007). "Structure based semantic measurement for information filtering agents." In *Proceedings of the Third Australasian Workshop on Advances in Ontologies-Volume 85* (pp. 25-33). Australian Computer Society, Inc..
<http://dl.acm.org/citation.cfm?id=2449262>
- Bobillo, Fernando, Miguel Delgado, Juan Gómez-Romero, and Enrique López (2009). "A semantic fuzzy expert system for a fuzzy balanced scorecard." *Expert Systems with Applications* 36(1): 423–433. <http://www.sciencedirect.com/science/article/pii/S0957417407004794>
- Bobillo, Fernando, and Umberto Straccia (2011). "Fuzzy ontology representation using OWL 2." *International Journal of Approximate Reasoning* 52(7): 1073–1094.
<http://www.sciencedirect.com/science/article/pii/S0888613X11000855>
- Bobrow, DG. (1964). "Natural language input for a computer problem solving system." <http://dspace.mit.edu/handle/1721.1/6903>
- Bock, J. (2008). "Parallel computation techniques for ontology reasoning." http://link.springer.com/chapter/10.1007/978-3-540-88564-1_60
- Boim, Rubi, Tova Milo, and Slava Novgorodov (2011). "DiRec: Diversified recommendations for semantic-less Collaborative Filtering." In *2011 IEEE 27th International Conference on Data Engineering*, IEEE, p. 1312–1315. <http://dl.acm.org/citation.cfm?id=2004686.2005671>
- Bojars, U, and JG Breslin (2008). "Interlinking the social web with semantics." *Intelligent Systems*, http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4525140
- Bojars, Uldis, John Breslin, and Stefan Decker (2008). "Social Networks and Data Portability using Semantic Web technologies." <http://aran.library.nuigalway.ie/xmlui/handle/10379/398>
-

-
- Borah, J. (2002). "Conceptual Modeling--The Missing Link of Simulation Development." In *Proceedings of the 2002 Spring Simulation Conference*.
- Borges, AM, M Corniel, and R Gil (2009). "Towards a study opportunities recommender system in ontological principles-based on semantic web environment." *WSEAS Transactions on Computers*, 8(2), 279-291. <http://dl.acm.org/citation.cfm?id=1512685>
- Borràs, Joan, Antonio Moreno, and Aida Valls (2014). "Intelligent tourism recommender systems: A survey." *Expert Systems with Applications* 41(16): 7370–7389. <http://www.sciencedirect.com/science/article/pii/S0957417414003431>
- Bosch, M. (2006). "Ontologies, different reasoning strategies, different logics, different kinds of knowledge representation: Working together." *Knowledge organization*. <http://eprints.rclis.org/archive/00009557>
- Bradley, P. (1999). "Intelligent Agents on the Web: An introduction to specific intelligent agents and a guide on how to familiarise yourself with applications to everyday life." *MANAGING INFORMATION-LONDON-ASSOCIATION OF INFORMATION MANAGEMENT*-, 35-41. http://scholar.google.es/scholar?q=Bradley,+Phil.+«Intelligent+agents+on+the+web».&btnG=&hl=es&as_sdt=0,5#0
- Brady, Erin, and Jeffrey P. Bigham (2014). "How companies engage customers around accessibility on social media." In *Proceedings of the 16th international ACM SIGACCESS conference on Computers & accessibility - ASSETS '14*, New York, New York, USA: ACM Press, p. 51–58. <http://dl.acm.org/citation.cfm?id=2661334.2661355>
- Bragaglia, S, F Chesani, and A Ciampolini (2010). "An hybrid architecture integrating forward rules with fuzzy ontological reasoning." In *Hybrid Artificial Intelligence Systems* (pp. 438-445). Springer Berlin Heidelberg. http://link.springer.com/chapter/10.1007/978-3-642-13769-3_53
- Breslin, J, and S Decker (2007). "The future of social networks on the internet: The need for semantics." *Internet Computing, IEEE*, 11(6), 86-90. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4376234
- Bridge, Derek, and Alex Ferguson (2002). "An Expressive Query Language for Product Recommender Systems." *Artificial Intelligence Review* 18(3-4): 269–307. <http://link.springer.com/article/10.1023/A:1020743321429>
- Brusilovsky, P. (2001). "Adaptive hypermedia." *User modeling and user-adapted interaction*. <http://link.springer.com/article/10.1023/A:1011143116306>
-

-
- Bu, J, S Tan, C Chen, C Wang, and H Wu (2010). "Music recommendation by unified hypergraph: combining social media information and music content." In *Proceedings of the international conference on Multimedia* (pp. 391-400). ACM. <http://dl.acm.org/citation.cfm?id=1874005>
- Bulte, C Van den, and S Wuyts (2007). "Social networks and marketing." *MSI Relevant Knowledge Series*.http://scholar.google.es/scholar?q=Social+networks+and+marketing&btnG=&hl=es&as_sdt=0,5#0
- Buono, P, MF Costabile, S Guida, and A Piccinno (2002). "Integrating user data and collaborative filtering in a web recommendation system." In *Hypermedia: openness, structural awareness, and adaptivity* (pp. 315-321). Springer Berlin Heidelberg.. http://link.springer.com/chapter/10.1007/3-540-45844-1_29
- Buriano, L., M. Marchetti, F. Carmagnola, F. Cena, C. Gena, and I. Torre (2006). "The Role of Ontologies in Context-Aware Recommender Systems." In *7th International Conference on Mobile Data Management (MDM'06)*, IEEE, p. 80-80. <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1630616>
- Burke, R. (2007). "Hybrid web recommender systems." In *The adaptive web* (pp. 377-408). Springer Berlin Heidelberg. http://link.springer.com/chapter/10.1007/978-3-540-72079-9_12
- Burke, RD, A Felfernig, and MH Göker (2011). "Recommender Systems: An Overview." *AI Magazine*. 32(3), 13-18. <http://gamejam.cti.depaul.edu/~rburke/pubs/burke-et-al-aimag11a.pdf>
- Burke, RD, KJ Hammond, and BC Yound (1997). "The FindMe approach to assisted browsing." *IEEE Expert*, 12(4), 32-40. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=608186
- Burke, RD, KJ Hammond, and BC Young (1996). "Knowledge-based navigation of complex information spaces." In *Proceedings of the national conference on artificial intelligence* (Vol. 462, p. 468). <http://www.aaai.org/Papers/AAAI/1996/AAAI96-069.pdf>
- Burke, Robin (2002). "Hybrid Recommender Systems: Survey and Experiments." *User Modeling and User-Adapted Interaction* 12(4): 331-370. <http://link.springer.com/article/10.1023/A:1021240730564>
- Cahn, S.M. (2012). "Classics of western philosophy." Hackett Publishing. <http://books.google.es/books?hl=es&lr=&id=y4j3--6rMcYC&oi=fnd&pg=PP1&dq=Classics+of+Western+Philosophy,+Sixth+ed.,+Hackett+Publishing+Company,+2002&ots=dR3aR2JEV&sig=fZWla8lpa2FSY8FQnFfXl1OerXk>
-

-
- Canberk, B, IF Akyildiz, and S Oktug (2011). "Primary user activity modeling using first-difference filter clustering and correlation in cognitive radio networks." *IEEE/ACM Transactions on Networking* <http://dl.acm.org/citation.cfm?id=1959455>
- Carbonaro, A, and R Ferrini (2007). "Ontology-based video annotation in multimedia entertainment." In *Proceedings of the 3rd IEEE International Workshop on Networking Issues in Multimedia Entertainment (NIME'07)-4th IEEE Communications and Networking Conference (CCNC 2007)*. <http://www.cs.unibo.it/~ferrini/papers/NIME2007.pdf>
- Cardoso, J. (2007). "The semantic web vision: Where are we?" *Intelligent Systems, IEEE*. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4338499
- Casado-Lumbreras, C. (2012). "PsyDis: Towards a diagnosis support system for psychological disorders." *Expert systems with applications*, 39(13), 11391-11403.. <http://www.sciencedirect.com/science/article/pii/S0957417412006306>
- Cea, G Aguado de, and E Montiel-Ponsoda (2012). "Term variants in ontologies." http://oa.upm.es/19602/1/INVE_MEM_2012_137765.pdf
- Chan, FTS (2005). "Application of a hybrid case-based reasoning approach in electroplating industry." *Expert Systems with Applications*, 29(1), 121-130. <http://www.sciencedirect.com/science/article/pii/S0957417405000126>
- Chang, E, TS Dillon, and F Hussain (2007). "Trust ontologies for e-service environments." *International Journal of Intelligent Systems*, 22(5), 519-545. <http://onlinelibrary.wiley.com/doi/10.1002/int.20212/abstract>
- Chen, Li, Wei Zeng, and Quan Yuan (2013). "A unified framework for recommending items, groups and friends in social media environment via mutual resource fusion." *Expert Systems with Applications* 40(8): 2889–2903. <http://www.sciencedirect.com/science/article/pii/S095741741201247X>
- Choi, Sang-Min, and Yo-Sub Han (2013). "Representative reviewers for Internet social media." *Expert Systems with Applications* 40(4): 1274–1282. <http://www.sciencedirect.com/science/article/pii/S095741741201024X>
- Christakou, Chistina, Spyros Vrettos, and Andreas Stafylopatis (2007). "A HYBRID MOVIE RECOMMENDER SYSTEM BASED ON NEURAL NETWORKS." *International Journal on Artificial Intelligence Tools* 16(05): 771–792. <http://www.worldscientific.com/doi/abs/10.1142/S0218213007003540>
-

-
- Christensen, Ingrid A., and Silvia Schiaffino (2011). "Entertainment recommender systems for group of users." *Expert Systems with Applications* 38(11): 14127–14135.
<http://www.sciencedirect.com/science/article/pii/S0957417411007482>
- Chu, H, and M Rosenthal (1996). "Search engines for the World Wide Web: A comparative study and evaluation methodology." *PROCEEDINGS OF THE ANNUAL MEETING-*
http://cui.unige.ch/tcs/cours/algoweb/2002/articles/art_habashi_arash.pdf
- Ciaramella, Alessandro, Mario G. C. A., Cimio, Beatrice Lazzerini, and Francesco Marcelloni (2010). "A SITUATION-AWARE RESOURCE RECOMMENDER BASED ON FUZZY AND SEMANTIC WEB RULES." *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 18(04): 411–430. <http://www.worldscientific.com/doi/abs/10.1142/S0218488510006623>
- Cintula, P, C Fermüller, and C Noguera (2015). "Handbook of Mathematical Fuzzy Logic-Volume 3." <http://philpapers.org/rec/CINHOM>
- Clark, Tim, and Marco Roos (2014). "Recent applications of web semantics in eLifeScience." *Web Semantics: Science, Services and Agents on the World Wide Web* 29: 1–2.
<http://www.sciencedirect.com/science/article/pii/S1570826814001085>
- Colombo-Mendoza, LO. (2015). "RecomMetz: A context-aware knowledge-based mobile recommender system for movie showtimes." *Expert Systems with Applications*, 42(3), 1202–1222. <http://www.sciencedirect.com/science/article/pii/S0957417414005557>
- Colomo-Palacios, R. (2012). "ReSySTER: A hybrid recommender system for Scrum team roles based on fuzzy and rough sets." *International Journal of Applied Mathematics and Computer Science*, 22(4), 801–816. <http://www.degruyter.com/view/j/amcs.2012.22.issue-4/v10006-012-0059-9/v10006-012-0059-9.xml>
- Copp, CB, and RL Ivy (2001). "Networking Trends of Small Tourism Businesses in Post-Socialist Slovakia." *Journal of Small Business Management*.
<http://onlinelibrary.wiley.com/doi/10.1111/0447-2778.00031/abstract>
- Correa, Teresa, Amber Willard Hinsley, and Homero Gil de Zúñiga (2010). "Who interacts on the Web?: The intersection of users' personality and social media use." *Computers in Human Behavior* 26(2): 247–253.
<http://www.sciencedirect.com/science/article/pii/S0747563209001472>
- Costantino, M. (1997). "Natural language processing and information extraction: Qualitative analysis of financial news articles." In *Computational Intelligence for Financial Engineering (CIFER)*,
-

-
- 1997., *Proceedings of the IEEE/IAFE* 1997 (pp. 116-122). IEEE..
http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=618923
- Crabtree, I Barry, and SJ Soltysiak (1998). "Identifying and tracking changing interests." *International Journal on Digital Libraries*. <http://www.springerlink.com/index/4H05HWWHG1YM60AQ.pdf>
- Cremonesi, Paolo, Antonio Tripodi, and Roberto Turrin. (2011). "Cross-Domain Recommender Systems." In *2011 IEEE 11th International Conference on Data Mining Workshops*, IEEE, p. 496–503. <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6137420>
- Cromity, Jamal (2012). "The Impact of Social Media in Review." *New Review of Information Networking* 17(1): 22–33. <http://dx.doi.org/10.1080/13614576.2012.673425>
- Cruz, RAPP, FJ García Peñalvo, and L Alonso Romero (2003). "Perfiles de usuario: En la senda de la personalización." <http://gredos.usal.es/jspui/handle/10366/21750>
- Cuadrado, JL López (2009). "Definición de un modelo de representación del conocimiento para procesos de estimación de presupuestos." <https://e-archivo.uc3m.es/handle/10016/6729>
- d'Aquin, Mathieu, Jean Lieber, and Amedeo Napoli (2006). Towards a semantic portal for oncology using a description logic with fuzzy concrete domains. *Capturing Intelligence*, 1, 379-393..
<http://www.sciencedirect.com/science/article/pii/S1574957606800213>
- Davies, J, and R Weeks (2004). "QuizRDF: Search technology for the semantic web." In *System Sciences, 2004. Proceedings of the 37th Annual Hawaii International Conference on* (pp. 8-pp). IEEE. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1265293
- Davis, R, H Shrobe, and P Szolovits (1993). "What is a knowledge representation?" *AI magazine*. 14 (1), 17. <http://www.aaai.org/ojs/index.php/aimagazine/article/viewArticle/1029>
- Decker, S, and S Melnik (2000). "The semantic web: The roles of XML and RDF." *Internet Computing IEEE*, 4(5), 63-73. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=877487
- Decotignie, JD, and P Pleinevaux (1993). "A survey on industrial communication networks." *Annales des télécommunications*. <http://link.springer.com/article/10.1007/BF02995472>
- Degemmis, Marco, Pasquale Lops, and Giovanni Semeraro (2007). "A content-collaborative recommender that exploits WordNet-based user profiles for neighborhood formation." *User Modeling and User-Adapted Interaction* 17(3): 217–255.
<http://link.springer.com/10.1007/s11257-006-9023-4>
-

-
- Dey, AK, GD Abowd, and D Salber (2001). "A conceptual framework and a toolkit for supporting the rapid prototyping of context-aware applications." *Human-computer interaction*. <http://dl.acm.org/citation.cfm?id=1463110>
- Díez, F.J., J. Mira, E. Iturralde, and S. Zubillaga (1997). "DIAVAL, a Bayesian expert system for echocardiography." *Artificial Intelligence in Medicine* 10(1): 59–73. <http://www.sciencedirect.com/science/article/pii/S0933365797003849>
- Drachsler, H, H Hummel, and R Koper (2007). "Recommendations for learners are different: Applying memory-based recommender system techniques to lifelong learning." <http://dspace.ou.nl/handle/1820/1018>
- Duan, Y., J.S. Edwards, and M.X. Xu (2005). "Web-based expert systems: benefits and challenges." *Information & Management* 42(6): 799–811. <http://www.sciencedirect.com/science/article/pii/S0378720604001193>
- Duffy, Peter D., and Axel Bruns (2006). "The Use of Blogs, Wikis and RSS in Education: A Conversation of Possibilities." <http://eprints.qut.edu.au/5398/1/5398.pdf>
- Dugan, Casey, Michael Muller, David R. Millen, Werner Geyer, Beth Brownholtz, and Marty Moore. (2007). "The dogear game." In *Proceedings of the 2007 international ACM conference on Conference on supporting group work - GROUP '07*, New York, New York, USA: ACM Press, p. 387. <http://dl.acm.org/citation.cfm?id=1316624.1316683>
- Eckert, W, E Levin, and R Pieraccini (1997). "User modeling for spoken dialogue system evaluation." In *Automatic Speech Recognition and Understanding, 1997. Proceedings., 1997 IEEE Workshop on* (pp. 80-87). IEEE. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=658991
- Eiter, T, G Ianni, T Krennwallner, and A Polleres (2008). "Rules and ontologies for the semantic web." In *Reasoning Web* (pp. 1-53). Springer Berlin Heidelberg. http://link.springer.com/chapter/10.1007/978-3-540-85658-0_1
- Enders, A, and H Hungenberg (2008). "The long tail of social networking.: Revenue models of social networking sites." *European Management Journal*, 26(3), 199-211.. <http://www.sciencedirect.com/science/article/pii/S0263237308000200>
- Eom, SB. (1996). "A survey of operational expert systems in business (1980–1993)." *Interfaces*. 26(5), 50-70. <http://pubsonline.informs.org/doi/abs/10.1287/inte.26.5.50>
- Felfernig, A, and G Friedrich (2006). "An integrated environment for the development of knowledge-based recommender applications." *International Journal of Electronic Commerce*, 11(2), 11-
-

-
34.
https://scholar.google.es/scholar?q=An+Integrated+Environment+for+the+Development+of+Knowledge-Based+Recommender+Applications&btnG=&hl=es&as_sdt=0%2C5#0
- Fensel, D. (2004). "Triple-space computing: Semantic Web Services based on persistent publication of information." *Intelligence in Communication Systems*.
http://link.springer.com/chapter/10.1007/978-3-540-30179-0_4
- Fensel, D. (2005). "Spinning the Semantic Web: bringing the World Wide Web to its full potential." Mit Press.
[http://books.google.es/books?hl=es&lr=&id=zQ34EoZO2IYC&oi=fnd&pg=PR11&dq=Berners-Lee,+T.,+\(2003\).+Foreword.+In+D.+Fensel+\(Ed.\),+Spinning+the+Semantic+Web:+Bringing+the+World+Wide+Web+to+its+full+potential+&ots=752s60CoJR&sig=yOhA5dyJGJxiuxoEwgTe3UUeF3Q](http://books.google.es/books?hl=es&lr=&id=zQ34EoZO2IYC&oi=fnd&pg=PR11&dq=Berners-Lee,+T.,+(2003).+Foreword.+In+D.+Fensel+(Ed.),+Spinning+the+Semantic+Web:+Bringing+the+World+Wide+Web+to+its+full+potential+&ots=752s60CoJR&sig=yOhA5dyJGJxiuxoEwgTe3UUeF3Q)
- Fensel, D, and F Van Harmelen (2001). "OIL: An ontology infrastructure for the semantic web." *IEEE intelligent systems*, (2), 38-45. <http://www.computer.org/csdl/mags/ex/2001/02/x2038.pdf>
- Fernández, Miriam, Iván Cantador, Vanesa López, David Vallet, Pablo Castells, and Enrico Motta (2011). "Semantically enhanced Information Retrieval: An ontology-based approach." *Web Semantics: Science, Services and Agents on the World Wide Web* 9(4): 434-452.
<http://www.sciencedirect.com/science/article/pii/S1570826810000910>
- Fernández-Tobías, Ignacio, Iván Cantador, Marius Kaminskas, and Francesco Ricci (2011). "A generic semantic-based framework for cross-domain recommendation." *Proceedings of the 2nd International Workshop on Information Heterogeneity and Fusion in Recommender Systems - HetRec '11*: 25-32. <http://dl.acm.org/citation.cfm?doid=2039320.2039324>.
- Foltz, PW, and ST Dumais (1992). "Personalized information delivery: An analysis of information filtering methods." *Communications of the ACM*. <http://dl.acm.org/citation.cfm?id=138866>
- Franke, Nikolaus, Martin Schreier, and Ulrike Kaiser (2010). "The 'I Designed It Myself' Effect in Mass Customization." *Management Science* 56(1): 125-140.
<http://pubsonline.informs.org/doi/abs/10.1287/mnsc.1090.1077>
- Frew, J., and R. Bose (2001). "Earth System Science Workbench: a data management infrastructure for earth science products." In *Proceedings Thirteenth International Conference on Scientific and Statistical Database Management. SSDBM 2001*, IEEE Comput. Soc, p. 180-189.
<http://ieeexplore.ieee.org/articleDetails.jsp?arnumber=938550>
-

-
- Friedman, C, and G Hripcsak (1999). "Natural language processing and its future in medicine." *Academic Medicine* 74(8), 890-5. http://journals.lww.com/academicmedicine/Abstract/1999/08000/Natural_language_processing_and_its_future_in.12.aspx
- Fundulaki, Irini, Josep Larriba Pey, David Dominguez-sal, Ioan Toma, Dieter Fensel, Barry Bishop, Thomas Neumann, Orri Erling, Peter Neubauer, Paul Groth, Frank Van Harmelen, and Peter Boncz (2012). "The Linked Data Benchmark Council (LDBC)." *Proceedings of the First European Data Forum* 877: 6–8. <http://ceur-ws.org/Vol-877/paper10.pdf>.
- Gandon, Fabien L, and Norman M Sadeh (2004). "Semantic web technologies to reconcile privacy and context awareness." *Web Semantics: Science, Services and Agents on the World Wide Web* 1(3): 241–260. <http://www.sciencedirect.com/science/article/pii/S1570826804000022>
- Garcia, Inma, Laura Sebastia, and Eva Onaindia (2011). "On the design of individual and group recommender systems for tourism." *Expert Systems with Applications* 38(6): 7683–7692. <http://www.sciencedirect.com/science/article/pii/S095741741001506X>
- Garcia-Crespo, A. (2010). "Conceptual model for semantic representation of industrial manufacturing processes." *Computers in Industry*, 61(7), 595-612.. <http://www.sciencedirect.com/science/article/pii/S0166361510000059>
- García-Crespo, Á, and A Rodríguez (2010). "ODDIN: Ontology-driven differential diagnosis based on logical inference and probabilistic refinements." *Expert Systems with Applications*, 37(3), 2621-2628. <http://www.sciencedirect.com/science/article/pii/S0957417409007702>
- García-Crespo, Ángel, José Luis López-Cuadrado, Ricardo Colomo-Palacios, Israel González-Carrasco, and Belén Ruiz-Mezcua (2011). "Sem-Fit: A semantic based expert system to provide recommendations in the tourism domain." *Expert Systems with Applications* 38(10): 13310–13319. <http://linkinghub.elsevier.com/retrieve/pii/S0957417411006798>
- García-Crespo, Ángel, José Luis López-Cuadrado, Israel González-Carrasco, Ricardo Colomo-Palacios, and Belén Ruiz-Mezcua (2012). "SINVLIO: Using semantics and fuzzy logic to provide individual investment portfolio recommendations." *Knowledge-Based Systems* 27: 103–118. <http://linkinghub.elsevier.com/retrieve/pii/S0950705111001912>
- Garton, L. (1997). "Studying online social networks." *Journal of Computer-Mediated Communication*, 3(1), 0-0. http://onlinelibrary.wiley.com/doi/10.1111/j.1083-6101.1997.tb00062.x/full?sms_ss=facebook&at_xt=4da62ce480237b90,0
-

-
- Gemünden, HG, T Ritter, and A Walter (1997). "Relationships and networks in international markets." <http://library.wur.nl/WebQuery/clc/981437>
- Genesereth, Michael R, and Richard E Fikes (1992). "Knowledge interchange format-version 3.0: Reference manual."
- Ghani, R, and A Fano (2002). "Building recommender systems using a knowledge base of product semantics." In *Proceedings of the Workshop on Recommendation and Personalization in ECommerce at the 2nd International Conference on Adaptive Hypermedia and Adaptive Web based Systems* (pp. 27-29). <http://images.accenture.ca/SiteCollectionDocuments/PDF/recommenderws02.pdf>
- Giarratano, J, and G Riley (1998). "Expert systems: principles and programming, 1998." *PWS-Kent, Boston, MA* PWS-Kent, .
http://scholar.google.es/scholar?q=Expert+Systems:+principles+and+programming,+&btnG=&hl=es&as_sdt=0,5#5
- Giarratano, JC, and G Riley (1989). "Expert systems: principles and programming." <http://dl.acm.org/citation.cfm?id=583478>
- Gilbert, E, and K Karahalios (2009). "Predicting tie strength with social media." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 211-220). ACM. <http://dl.acm.org/citation.cfm?id=1518736>
- Ginsberg, ML, and DF Geddis (1991). "Is there any need for domain-dependent control information?" In *AAAI* (pp. 452-457). AAAI. http://scholar.google.es/scholar?q=Ginsberg,+1991&btnG=&hl=es&as_sdt=0,5#1
- Glass, J, S Marx, T Schmidt, and F Sivrikaya (2010). "Semantic TV Engine: An IPTV Enabler for Personalized Recommendations." *Information Management SPIM* 2010, 34. <http://hmk.ffzg.hr/bibl/lrec2010/workshops/W7.pdf#page=38>
- Glorot, X, A Bordes, and Y Bengio (2011). "Domain adaptation for large-scale sentiment classification: A deep learning approach." In *Proceedings of the 28th International Conference on Machine Learning (ICML-11)* (pp. 513-520). http://machinelearning.wustl.edu/mlpapers/paper_files/ICML2011Glorot_342.pdf
- Golbeck, J. (2006). "Combining provenance with trust in social networks for semantic web content filtering." *Provenance and Annotation of Data*. http://link.springer.com/chapter/10.1007/11890850_12
-

-
- Golbeck, J, and J Hendler (2006). "Filmtrust: Movie recommendations using trust in web-based social networks." In *Proceedings of the IEEE Consumer communications and networking conference* (Vol. 96, pp. 282-286). <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.74.913&rep=rep1&type=pdf>
- Golbeck, J, and A Mannes (2006). "Using Trust and Provenance for Content Filtering on the Semantic Web." In *MTW*. http://www.ra.ethz.ch/CDstore/www2006/www.l3s.de/~olmedilla/events/MTW06_papers/paper29.pdf
- Goldberg, David, David Nichols, Brian M. Oki, and Douglas Terry (1992). "Using collaborative filtering to weave an information tapestry." *Communications of the ACM* 35(12): 61–70. <http://dl.acm.org/citation.cfm?id=138859.138867>
- Gomez-Perez, A., and O. Corcho (2002). "Ontology languages for the Semantic Web." *IEEE Intelligent Systems* 17(1): 54–60. <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=988453>
- Gonzalez, AJ, and DD Dankel (1993). "The engineering of knowledge-based systems: theory and practice." <http://dialnet.unirioja.es/servlet/libro?codigo=370879>
- Gonzalez, Avelino J, and Douglas D Dankel (1993). *The engineering of knowledge-based systems: theory and practice*. ed. Englewood Cliffs. New Jersey, USA: Prentice hall Englewood Cliffs (NJ).
- Gonzalez-Carrasco, I. (2012). "PB-ADVISOR: A private banking multi-investment portfolio advisor." *Information Sciences*, 206, 63-82. <http://www.sciencedirect.com/science/article/pii/S0020025512002526>
- Gruber, Thomas R. (1992). *Ontolingua: A mechanism to support portable ontologies*. Stanford University, Knowledge Systems Laboratory.
- Gruber, T. R. (1993). "A translation approach to portable ontology specifications." *Knowledge Acquisition* 5(2): 199–220. <http://www.sciencedirect.com/science/article/pii/S1042814383710083>.
- Guarino, N. (1998). "Formal ontology in information systems: Proceedings of the first international conference (FOIS'98), June 6-8, Trento, Italy." [http://books.google.es/books?hl=es&lr=&id=Wf5p3_fUxacC&oi=fnd&pg=PR5&dq=Formal+Ontology+and+Information+Systems".+Proceedings+of+FOIS+'98&ots=nmRJXZqCLL&sig=aJH-fXdFmXQEJyR9e2uq7xKcGY](http://books.google.es/books?hl=es&lr=&id=Wf5p3_fUxacC&oi=fnd&pg=PR5&dq=Formal+Ontology+and+Information+Systems)
-

-
- Guha, R., and Rob McCool (2003). "TAP: a Semantic Web platform." *Computer Networks* 42(5): 557–577. <http://www.sciencedirect.com/science/article/pii/S1389128603002251>
- Guo, Y, and J Heflin (2006). "A scalable approach for partitioning OWL knowledge bases." In *Proceedings of the 2nd International Workshop on Scalable Semantic Web Knowledge Base Systems* (pp. 636-641). <http://swat.cse.lehigh.edu/pubs/guo06b.pdf>
- Guy, Ido, Naama Zwerdling, David Carmel, Inbal Ronen, Erel Uziel, Sivan Yogev, and Shila Ofek-Koifman (2009). "Personalized recommendation of social software items based on social relations." In *Proceedings of the third ACM conference on Recommender systems - RecSys '09*, New York, New York, USA: ACM Press, p. 53. <http://dl.acm.org/citation.cfm?id=1639714.1639725>
- Guy, Ido, Naama Zwerdling, Inbal Ronen, David Carmel, and Erel Uziel (2010). "Social media recommendation based on people and tags." In *Proceeding of the 33rd international ACM SIGIR conference on Research and development in information retrieval - SIGIR '10*, New York, New York, USA: ACM Press, p. 194. <http://dl.acm.org/citation.cfm?id=1835449.1835484>
- Ha, Inay, Kyeong-jin Oh, Myung-duk Hong, and Geun-sik Jo (2012). "LNAI 7653 - Social Filtering Using Social Relationship for Movie Recommendation." : 395–404.
- Hadavandi, E, H Shavandi, and A Ghanbari (2010). "Integration of genetic fuzzy systems and artificial neural networks for stock price forecasting." *Knowledge-Based Systems*. <http://www.sciencedirect.com/science/article/pii/S0950705110000857>
- Hassanein, Khaled, and Milena Head (2014). "The Impact of Infusing Social Presence in the Web Interface: An Investigation Across Product Types." http://www.tandfonline.com/doi/abs/10.2753/JEC1086-4415100202#.VSGvV_msWSO
- Hendler, J. (2009). "Web 3.0 Emerging." *Computer* 42(1), 111-113.. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4755170
- Hendler, James A. (2005). "Knowledge Is Power: A View from the Semantic Web." *AI Magazine* 26(4): 76. <http://www.aaai.org/ojs/index.php/aimagazine/article/view/1851>
- Hepp, M. (2008). "Goodrelations: An ontology for describing products and services offers on the web." *Knowledge Engineering: Practice and Patterns*. http://link.springer.com/chapter/10.1007/978-3-540-87696-0_29
- Herlocker, JL, and JA Konstan (2004). "Evaluating collaborative filtering recommender systems." *ACM Transactions on Information Systems (TOIS)*, 22(1), 5-53.. <http://dl.acm.org/citation.cfm?id=963772>
-

-
- Hernández del Olmo, Félix, and Elena Gaudioso (2008). "Evaluation of recommender systems: A new approach." *Expert Systems with Applications* 35(3): 790–804. <http://www.sciencedirect.com/science/article/pii/S0957417407002928>
- Hoekstra, Rinke (2009). "Ontology Representation Design Patterns and Ontologies that Make Sense." : 1–236. <http://dl.acm.org/citation.cfm?id=1659447.1659449>
- Hofmann, Thomas (2004). "Latent semantic models for collaborative filtering." *ACM Transactions on Information Systems* 22(1): 89–115. <http://dl.acm.org/citation.cfm?id=963770.963774>
- Horridge, M, and PF Patel-Schneider (2009). "OWL 2 web ontology language manchester syntax." *W3C Working Group Note*. <http://www.w3.org/2009/pdf/NOTE-owl2-manchester-syntax-20091027.pdf>
- Horrocks, I, PF Patel-Schneider, and F van Harmelen (2002). "Reviewing the design of DAML+ OIL: An ontology language for the semantic web." *AAAI/IAAI*. <http://www.aaai.org/Papers/AAAI/2002/AAAI02-119.pdf?pagewanted=all>
- Hsieh, Kun-Lin (2011). "Employing a recommendation expert system based on mental accounting and artificial neural networks into mining business intelligence for study abroad's P/S recommendations." *Expert Systems with Applications* 38(12): 14376–14381. <http://www.sciencedirect.com/science/article/pii/S0957417411004866>
- Huang, Z, W Chung, TH Ong, and H Chen (2002). "A graph-based recommender system for digital library." In *Proceedings of the 2nd ACM/IEEE-CS joint conference on Digital libraries* (pp. 65–73). ACM. <http://dl.acm.org/citation.cfm?id=544231>
- Huming, Gao, and Li Weili (2010). "A Hotel Recommendation System Based on Collaborative Filtering and Rankboost Algorithm." *2010 Second International Conference on Multimedia and Information Technology*: 317–320. <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5474286>
- Hung-Wen Tung, and Von-Wun Soo (2004). "A personalized restaurant recommender agent for mobile e-service." In *IEEE International Conference on e-Technology, e-Commerce and e-Service, 2004. EEE '04. 2004*, IEEE, p. 259–262. <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1287319>
- Hunter, R, J Perez, J Marquez, and JC Hernandez (2007). "Modeling the integration between technological product specifications and inspection process." *Journal of materials processing technology*, 191(1), 34–38. <http://www.sciencedirect.com/science/article/pii/S0924013607002324>
-

-
- Hyung, Ziwon, Kibeom Lee, and Kyogu Lee (2013). "Music recommendation using text analysis on song requests to radio stations." *Expert Systems with Applications*.
<http://www.sciencedirect.com/science/article/pii/S095741741300852X>
- J, K GEORGE, and Y Bo (2008). "Fuzzy sets and fuzzy logic, theory and applications." -.
<http://digilib.uin-suka.ac.id/7049/>
- Jannach, D, M Zanker, A Felfernig, and G Friedrich (2010). "Recommender systems: an introduction." Cambridge University Press.
http://books.google.es/books?hl=es&lr=&id=eygTJBd_U2cC&oi=fnd&pg=PR5&dq=related:rUCjhLcnoXUJ:scholar.google.com/&ots=mUu571ySwK&sig=DZB_Ignw1uzBfJDfOUV3Uj6qypc
- Jeong, D, M Choi, YS Jeon, and YH Han (2006). "A novel memory-oriented OWL storage system." *Frontiers of High* http://link.springer.com/chapter/10.1007/11942634_55
- Jr, A Rabelo, AR Rocha, and K Oliveira (1997). "An expert system for diagnosis of acute myocardial infarction with ECG analysis." *Artificial Intelligence in Medicine*, 10(1), 75-92.
<http://www.sciencedirect.com/science/article/pii/S0933365797003850>
- Jung, JJ, and J Euzenat (2007). "Towards semantic social networks." *The semantic web: research and applications*. http://link.springer.com/chapter/10.1007/978-3-540-72667-8_20
- Kahney, L. (2003). "Making friendsters in high places." *Wired*, July.
[http://scholar.google.es/scholar?q=Making+Friendsters+in+High+Places+Wired+\(July+2003\)&btnG=&hl=es&as_sdt=0,5#1](http://scholar.google.es/scholar?q=Making+Friendsters+in+High+Places+Wired+(July+2003)&btnG=&hl=es&as_sdt=0,5#1)
- Kaplan, Andreas M., and Michael Haenlein (2009). "The fairyland of Second Life: Virtual social worlds and how to use them." *Business Horizons* 52(6): 563-572.
<http://www.sciencedirect.com/science/article/pii/S0007681309000895>
- Kaplan, A. M., & Haenlein, M. (2010). "Users of the world, unite! The challenges and opportunities of Social Media." *Business Horizons* 53(1): 59-68.
<http://www.sciencedirect.com/science/article/pii/S0007681309001232>
- Kaplan, A. M., & Haenlein, M. (2011). "The early bird catches the news: Nine things you should know about micro-blogging." *Business Horizons* 54(2): 105-113.
<http://www.sciencedirect.com/science/article/pii/S0007681310001254>
- Kherfi, ML, D Ziou, and A Bernardi (2004). "Image retrieval from the world wide web: Issues, techniques, and systems." *ACM Computing Surveys (CSUR)*.
<http://dl.acm.org/citation.cfm?id=1013210>
-

-
- Kietzmann, Jan H., Kristopher Hermkens, Ian P. McCarthy, and Bruno S. Silvestre (2011). "Social media? Get serious! Understanding the functional building blocks of social media." *Business Horizons* 54(3): 241–251.
<http://www.sciencedirect.com/science/article/pii/S0007681311000061>
- Kimball, L, and H Rheingold (2000). "How online social networks benefit organizations." *Rheingold Associates*. <https://groupjazz.com/pdf/osn.pdf>
- Kindberg, T, J Barton, and J Morgan (2002). "People, places, things: Web presence for the real world." *Mobile Networks and* <http://dl.acm.org/citation.cfm?id=582470>
- Kinsella, Sheila, Alexandre Passant, John G. Breslin, Stefan Decker, and Ajit Jaokar (2009). *Advances in Computers Social Networking and The Web*. Elsevier.
<http://www.sciencedirect.com/science/article/pii/S0065245809010043>
- Kiryakov, Atanas, Borislav Popov, Ivan Terziev, Dimitar Manov, and Damyan Ognyanoff (2011). "Semantic Annotation, Indexing, and Retrieval." *Web Semantics: Science, Services and Agents on the World Wide Web* 2(1). <http://imap.websemanticsjournal.org/index.php/ps/article/view/53>
- Klein, GA, and JE Orasanu (1993). "Decision making in action: Models and methods." *This book is an* <http://psycnet.apa.org/psycinfo/1993-97634-000>
- Klein, M. (2002). "Interpreting XML documents via an RDF schema ontology." In *Proceedings. 13th International Workshop on Database and Expert Systems Applications*, IEEE Comput. Soc, p. 889–893. <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1046008>
- Klein, MCA, and D Fensel (2001). "Ontology versioning on the Semantic Web." *SWWS*. <http://secs.ceas.uc.edu/~mazlack/ECE.716.Sp2011/Semantic.Web.Ontology.Papers/klein01ontology.pdf>
- Kobeissy, N. (2009). "Towards a context management framework: context description, modelling, and collect in the domain of mobile networks." <http://www.theses.fr/2009TELE0005>
- Kolodner, Janet L. (1992). "An introduction to case-based reasoning." *Artificial Intelligence Review* 6(1): 3–34. <http://link.springer.com/10.1007/BF00155578>
- Konstan, Joseph A. (2004). "Introduction to recommender systems" ed. Joseph A. Konstan. *ACM Transactions on Information Systems* 22(1): 1–4.
<http://dl.acm.org/citation.cfm?id=963770.963771>
- Koren, Y, R Bell, and C Volinsky (2009). "Matrix factorization techniques for recommender systems." *Computer*. <http://pascal.computer.org/csdl/mags/co/2009/08/mco2009080030.html>
-

-
- Kosak, DM, and AK Lang (1999). "System and method employing individual user content-based data and user collaborative feedback data to evaluate the content of an information entity in a large." *US Patent 5,983,214*. <http://www.google.com/patents/US5983214>
- Kotler, P, and D Gertner (2002). "Country as brand, product, and beyond: A place marketing and brand management perspective." *The Journal of Brand Management*. <http://www.ingentaconnect.com/content/pal/bm/2002/00000009/00000004/art00003>
- Kwak, Haewoon, Changhyun Lee, Hosung Park, and Sue Moon (2010). "What is Twitter, a social network or a news media?" In *Proceedings of the 19th international conference on World wide web - WWW '10*, New York, New York, USA: ACM Press, p. 591. <http://dl.acm.org/citation.cfm?id=1772690.1772751>
- Lame, G. (2005). "Using NLP techniques to identify legal ontology components: concepts and relations." *Law and the Semantic Web*. http://link.springer.com/chapter/10.1007/978-3-540-32253-5_11
- Lang, K. (1995). "Newsweeder: Learning to filter netnews." In *Twelfth International Conference on Machine Learning*. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.85.7363>
- Lassila, O, and F van Harmelen (2000). "The semantic web and its languages." *Intelligent Systems and their Applications, IEEE*, 15(6), 67-73. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=895864
- Le, MT, NT Nguyen, and J Swiatek (2010). "Intelligent Information and Database Systems: Second International Conference, ACIIDS 2010, Hue City, Vietnam, March 24-26, 2010, Proceedings (Vol. 5991). Springer. [https://books.google.es/books?hl=es&lr=&id=RtYiHnmXZq0C&oi=fnd&pg=PR2&dq=Standard+additive+fuzzy+system+for+stock+price+forecasting,+in+N.T.+Nguyen,+M.T.+Le+and+J.+Swiatek+\(Eds.\),+Proceedings+of+the+2nd+International+Conference+on+Intelligent+Information+and+Data&ots=Lo41rFqJfN&sig=cO9oukbsn8h6QARoDgMmlqHbneQ](https://books.google.es/books?hl=es&lr=&id=RtYiHnmXZq0C&oi=fnd&pg=PR2&dq=Standard+additive+fuzzy+system+for+stock+price+forecasting,+in+N.T.+Nguyen,+M.T.+Le+and+J.+Swiatek+(Eds.),+Proceedings+of+the+2nd+International+Conference+on+Intelligent+Information+and+Data&ots=Lo41rFqJfN&sig=cO9oukbsn8h6QARoDgMmlqHbneQ)
- Lee, C.-S., Z.-W. Jian, and L.-K. Huang (2005). "A Fuzzy Ontology and Its Application to News Summarization." *IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)* 35(5): 859–880. <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1510764>
- Lee, Wei-Po (2004). "Towards agent-based decision making in the electronic marketplace: interactive recommendation and automated negotiation." *Expert Systems with Applications* 27(4): 665–679. <http://www.sciencedirect.com/science/article/pii/S0957417404000739>
-

-
- Lehmann, Thomas M, Mark O Güld, Thomas Deselaers, Daniel Keysers, Henning Schubert, Klaus Spitzer, Hermann Ney, and Berthold B Wein (2005). "Automatic categorization of medical images for content-based retrieval and data mining." In *Computerized medical imaging and graphics: the official journal of the Computerized Medical Imaging Society* 29(2-3): 143–55. <http://www.sciencedirect.com/science/article/pii/S0895611104001168>
- Leis, Ángela, Miguel Ángel Mayer, Javier Torres Niño, Alejandro Rodríguez-González, Josep Maria Suelves, and Manuel Armayones (2013). "[Healthy eating support groups on Facebook: content and features]." *Gaceta sanitaria / S.E.S.P.A.S* 27(4): 355–7. <http://www.sciencedirect.com/science/article/pii/S0213911112003469>
- Leite Dantas Bezerra, Byron, and Francisco de Assis Tenorio de Carvalho (2010). "Symbolic data analysis tools for recommendation systems." *Knowledge and Information Systems* 26(3): 385–418. <http://link.springer.com/10.1007/s10115-009-0282-3>
- Lemos, Miguel Lagares, Daniel Villanueva Vasquez, Mateusz Radzinski, Lagares Lemos, and Juan Miguel Gómez-berbís (2012). "RING : A Context Ontology for Communication Channel Rule-based Recommender System 1 Introduction." In , p. 10. <http://sisinflab.poliba.it/sersy2012/>.
- Lerman, K. (2006). "Social networks and social information filtering on digg." *arXiv preprint cs/0612046*. <http://arxiv.org/abs/cs/0612046>
- Leung, Daniel, Rob Law, Hubert van Hoof, and Dimitrios Buhalis (2013). "Social Media in Tourism and Hospitality: A Literature Review." *Journal of Travel & Tourism Marketing* 30(1-2): 3–22. <http://dx.doi.org/10.1080/10548408.2013.750919>
- Li, C, and F Chan (2011). "Complex-Fuzzy Adaptive Image Restoration–An Artificial-Bee-Colony-Based Learning Approach." *Intelligent Information and Database Systems*. http://link.springer.com/chapter/10.1007/978-3-642-20042-7_10
- Li, J, Y Ding, Y Shi, and J Zhang (2010). "Building a Large Annotation Ontology for Movie Video Retrieval." *JDCTA*. <http://www.aicit.org/JDCTA/ppl/08.JDCTA2-390081.pdf>
- Li, L, and I Horrocks (2004). "A software framework for matchmaking based on semantic web technology." *International Journal of Electronic Commerce*. <http://mesharpe.metapress.com/index/6C3XD67PPYMBK8L9.pdf>
- Liang Hu, Jian Cao, Guandong Xu, Longbing Cao, Zhhiping Gu, Can Zhu (2013). "Personalized Recommendation via." : 595–605.
- Liddy, Elizabeth (2001). "Natural Language Processing." *The School of Information Studies Faculty Scholarship*. <http://surface.syr.edu/istpub/63>
-

-
- Liu, B. (2011). "Opinion mining and sentiment analysis." *Web Data Mining*. http://link.springer.com/chapter/10.1007/978-3-642-19460-3_11
- Liu, J, P Dolan, and ER Pedersen (2010). "Personalized news recommendation based on click behavior." In *Proceedings of the 15th international conference on Intelligent user interfaces* (pp. 31-40). ACM. <http://dl.acm.org/citation.cfm?id=1719976>
- Liu, J, W Wang, F Golnaraghi, and E Kubica (2010). "A neural fuzzy framework for system mapping applications." *Knowledge-Based Systems*. 23(6), 572-579. <http://www.sciencedirect.com/science/article/pii/S0950705110000596>
- Lopez-Cuadrado, JL. (2012). "CoKIM: Collaborative and Social Knowledge-Based Incident Manager." In *Proceedings of the 2012 International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2012)* (pp. 1211-1214). IEEE Computer Society. <http://dl.acm.org/citation.cfm?id=2456902>
- López-Cuadrado, JL. (2012). "SABUMO: Towards a collaborative and semantic framework for knowledge sharing." *Expert Systems with Applications*, 39(10), 8671-8680..... <http://www.sciencedirect.com/science/article/pii/S0957417412002266>
- Lops, P, M de Gemmis, and G Semeraro (2009). "A semantic content-based recommender system integrating folksonomies for personalized access." *Web Personalization in Intelligent Environments* (pp. 27-47). Springer Berlin Heidelberg. http://link.springer.com/chapter/10.1007/978-3-642-02794-9_2
- Lops Pasquale, Marco de Gemmis and Giovanni Semeraro (2011). "Content-based Recommender Systems: State of the Art and Trends." In *Recommender systems handbook* (pp. 73-105). Springer US. <https://www.google.es/search?q=Content-based+Recommender+Systems:+State+of+the+Art+and+Trends&oq=Content-based+Recommender+Systems:+State+of+the+Art+and+Trends&aqs=chrome.69i57j69i61j69i62l2.231j0&sourceid=chrome&ie=UTF-8>.
- Lovejoy, K, RD Waters, and GD Saxton (2012). "Engaging stakeholders through Twitter: How nonprofit organizations are getting more out of 140 characters or less." *Public Relations Review*. <http://www.sciencedirect.com/science/article/pii/S0363811112000082>
- Lu, J, D Wu, M Mao, W Wang, and G Zhang (2015). "Recommender System Application Developments: A Survey." *Decision Support Systems*. <http://www.sciencedirect.com/science/article/pii/S0167923615000627>
-

-
- Lü, Linyuan, Matúš Medo, Chi Ho Yeung, Yi-Cheng Zhang, Zi-Ke Zhang, and Tao Zhou (2012). "Recommender systems." *Physics Reports* 519(1): 1–49. <http://linkinghub.elsevier.com/retrieve/pii/S0370157312000828>
- Ma, Z.M., Fu Zhang, Li Yan, and Jingwei Cheng (2011). "Representing and reasoning on fuzzy UML models: A description logic approach." *Expert Systems with Applications* 38(3): 2536–2549. <http://www.sciencedirect.com/science/article/pii/S0957417410008389>
- MacGregor, R, and R Bates. (1987). "The Loom Knowledge Representation Language." <http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA183415>
- Maedche, A, B Motik, and L Stojanovic (2003). "Managing multiple and distributed ontologies on the Semantic Web." *The VLDB Journal—The International Journal on Very Large Data Bases*, 12(4), 286–302. <http://dl.acm.org/citation.cfm?id=953240>
- Maedche, A., B. Motik, and L. Stojanovic (2003). "Managing multiple and distributed ontologies on the Semantic Web." *The VLDB Journal The International Journal on Very Large Data Bases* 12(4): 286–302. <http://dl.acm.org/citation.cfm?id=953238.953240>
- Maedche, A., and S. Staab (2001). "Ontology learning for the Semantic Web." *IEEE Intelligent Systems* 16(2): 72–79. <http://www.computer.org/csdl/mags/ex/2001/02/x2072.html>
- Mair, A. (2004). "Recommender Systems for Tourism." *Advanced Interface Design*. https://personal.cis.strath.ac.uk/mark.dunlop/teaching/aid/AIDProceedings2006_v2.pdf#page=16
- Mak, Harry, Irena Koprinska, and Josiah Poon (2003). "INTIMATE: A Web-Based Movie Recommender Using Text Categorization." : 602. <http://dl.acm.org/citation.cfm?id=946251.947114>
- Mamdani, E.H., and S. Assilian (1975). "An experiment in linguistic synthesis with a fuzzy logic controller." *International Journal of Man-Machine Studies* 7(1): 1–13. <http://www.sciencedirect.com/science/article/pii/S0020737375800022>
- Mamdani, EH (1988). "Fuzzy sets and applications: selected papers by LA Zadeh: RR Yager, S Ovchinnikov, RM Tong, HT Nguyen (eds) John Wiley and Sons Inc, £ 45.85, ISBN 0." *Knowledge-Based Systems*. https://scholar.google.es/scholar?q=+Fuzzy+sets+and+applications%3A+selected+papers+by+L.+A.+Zadeh%3A+R+Yager%2C+S+Ovchinnikov%2C+R+M+Tong%2C+H+T+Nguyen+%28eds%29.+Knowledge-Based+Systems%2C+&btnG=&hl=es&as_sdt=0%2C5#1
-

-
- Mangold, W. Glynn, and David J. Faulds (2009). "Social media: The new hybrid element of the promotion mix." *Business Horizons* 52(4): 357–365.
<http://www.sciencedirect.com/science/article/pii/S0007681309000329>
- Martinez, Ana, Jose Arias, Ana Vilas, Jorge Garcia Duque, and Martin Lopez Nores (2009). "What's on TV tonight? An efficient and effective personalized recommender system of TV programs." *IEEE Transactions on Consumer Electronics* 55(1): 286–294.
<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4814447>
- Martínez, Luis, Manuel J. Barranco, Luis G. Pérez, and Macarena Espinilla (2008). "A Knowledge Based Recommender System with Multigranular Linguistic Information." *International Journal of Computational Intelligence Systems* 1(3): 225–236.
<http://dx.doi.org/10.1080/18756891.2008.9727620>
- Massa, P, and P Avesani (2004). "Trust-aware collaborative filtering for recommender systems." On the Move to Meaningful Internet Systems 2004: CoopIS, DOA, and ODBASE. Springer.
http://link.springer.com/chapter/10.1007/978-3-540-30468-5_31
- Masters, J, T Madhyastha, and A Shakouri (2008). "ExplaNet: A collaborative learning tool and hybrid recommender system for student-authored explanations." *Journal of Interactive Learning Research*, 19(1), 51-74. <http://www.editlib.org/p/21960/>
- McCarthy, J.F. (2001). "The virtual world gets physical: perspectives on personalization." *IEEE Internet Computing* 5(6): 48–53.
<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=968831>
- McCarthy, JF. (2002). "Pocket restaurantfinder: A situated recommender system for groups." In *Workshop on Mobile Ad-Hoc Communication at the 2002 ACM Conference on Human Factors in Computer Systems*.
<http://interrelativity.com/joe/publications/PocketRestaurantFinder-CHI2002ws-AdHoc.pdf>
- Mcguinness, Deborah L, Richard Fikes, Lynn Andrea Stein, and Franklin W Olin (2002). "DAML + OIL : An Ontology Language for the Semantic Web."
- Mcsherry, David (2005). "Explanation in Recommender Systems." *Artificial Intelligence Review* 24(2): 179–197. <http://link.springer.com/10.1007/s10462-005-4612-x>
- Michaelidou, Nina, Nikoletta Theofania Siamagka, and George Christodoulides (2011). "Usage, barriers and measurement of social media marketing: An exploratory investigation of small and medium B2B brands." *Industrial Marketing Management* 40(7): 1153–1159.
<http://www.sciencedirect.com/science/article/pii/S0019850111001374>
-

-
- Mika, Peter (2005). "Flink: Semantic Web technology for the extraction and analysis of social networks." *Web Semantics: Science, Services and Agents on the World Wide Web* 3(2-3): 211–223. <http://www.sciencedirect.com/science/article/pii/S1570826805000089>
- Milicevic, Aleksandra Klasnja, Alexandros Nanopoulos, and Mirjana Ivanovic (2010). "Social tagging in recommender systems: a survey of the state-of-the-art and possible extensions." *Artificial Intelligence Review* 33(3): 187–209. <http://link.springer.com/10.1007/s10462-009-9153-2>
- Miller, BN, I Albert, and SK Lam (2003). "MovieLens unplugged: experiences with an occasionally connected recommender system." In *Proceedings of the 8th international conference on Intelligent user interfaces* (pp. 263-266). ACM. <http://dl.acm.org/citation.cfm?id=604094>
- Mizoguchi, R, and J Bourdeau (2000). "Using ontological engineering to overcome common AI-ED problems." *Journal of Artificial Intelligence and Education*, 11, 107-121. <http://telearn.archives-ouvertes.fr/hal-00190087/>
- Mladenec, D. (1999). "Text-learning and related intelligent agents: a survey." *IEEE Intelligent Systems* 14(4): 44–54. <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=784084>
- Mobasher, B, H Dai, T Luo, and M Nakagawa (2002). "Discovery and evaluation of aggregate usage profiles for web personalization." *Data mining and knowledge discovery*, 6(1), 61-82. <http://link.springer.com/article/10.1023/A:1013232803866>
- Mobasher, B, X Jin, and Y Zhou (2004). "Semantically enhanced collaborative filtering on the web." *Web Mining: From Web to Semantic Web*. http://link.springer.com/chapter/10.1007/978-3-540-30123-3_4
- Monachesi, P, and T Markus (2010). "Using social media for ontology enrichment." *The Semantic Web: Research and Applications*. http://link.springer.com/chapter/10.1007/978-3-642-13489-0_12
- Monfil-Contreras, E. U., Alor-Hernandez, G., Cortes-Robles, G., Rodriguez-Gonzalez, A., & Hernandez-Chan, G. (2012). "A Survey of Techniques for Developing Recommendation Systems." In *International Conference on Electronics, Information and Communication Engineering* (EICE 2012). ASME Press. <http://ebooks.asmedigitalcollection.asme.org/content.aspx?bookid=408§ionid=38787962>
- Montaner, Miquel (2003). "A Taxonomy of Recommender Agents on the Internet." : 285–330.
- Morstatter, F, J Pfeffer, H Liu, and KM Carley (2013). "Is the sample good enough? Comparing data from Twitter's streaming API with Twitter's firehose." *Proceedings of ICWSM*. <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM13/paper/viewPDFInterstitial/6071/6379>
-

-
- Motta, Enrico (1998). "An overview of the OCML modelling language." In *the 8th Workshop on Methods and Languages*.
- Muntjewerff, A, and B Bredeweg (1999). "Ontological modelling for designing educational systems." <http://staff.science.uva.nl/~bouwer/pdf/ws-ai99.pdf>
- Murugesan, S. (2007). "Understanding Web 2.0." *IT professional*. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4287373
- Musen, MA. (1989). "Conceptual models of interactive knowledge acquisition tools." *Knowledge Acquisition*. <http://www.sciencedirect.com/science/article/pii/S1042814389800068>
- Nallasivan, S, and DD Price (2012). "Artificial intelligence for social media." *US Patent 8,249,996*. <http://www.google.com/patents/US8249996>
- Negnevitsky, M. (2005). "Artificial intelligence: a guide to intelligent systems." <https://books.google.es/books?hl=es&lr=&id=1BxYQnrfv9MC&oi=fnd&pg=PR11&dq=A+Guide+to+Intelligent+Systems,+&ots=GV9vZKRhZv&sig=bX5grcHUNRt7TEvRfrFjh745z84>
- Ngai, EWT, and FKT Wat. (2005). "Fuzzy decision support system for risk analysis in e-commerce development." *Decision support systems*. <http://www.sciencedirect.com/science/article/pii/S0167923604000272>
- Nilsson, A, and R Muradore (2009). "Ontology for robotics: A roadmap." In *14th International Conference on Advanced Robotics* (pp. 291-296). IEEE. <http://lup.lub.lu.se/record/1507114>
- Di Noia, Tommaso, Roberto Mirizzi, Vito Claudio Ostuni, Davide Romito, and Markus Zanker (2012). "Linked open data to support content-based recommender systems." In *Proceedings of the 8th International Conference on Semantic Systems - I-SEMANTICS '12*, New York, New York, USA: ACM Press, p. 1. <http://dl.acm.org/citation.cfm?id=2362499.2362501>
- Nosek, JT, and I Roth (1990). "A comparison of formal knowledge representationschemes as communication tools: predicate logic< i> vs</i> semantic network." *International Journal of Man-Machine Studies*. <http://www.sciencedirect.com/science/article/pii/S0020737305800306>
- Noy, NF, and M Klein (2004). "Ontology evolution: Not the same as schema evolution." *Knowledge and information systems*. <http://link.springer.com/article/10.1007/s10115-003-0137-2>
- Noy, NF, M Sintek, and S Decker (2001). "Creating semantic web contents with protege-2000." *IEEE intelligent systems*, (2), 60-71. <http://www.computer.org/csdl/mags/ex/2001/02/x2060.pdf>
-

-
- O'Keefe, RM, O Balci, and EP Smith (1986). "Validation of expert system performance." <http://dl.acm.org/citation.cfm?id=903416>
- Obitko, M. (2001). "Ontologies--Description and Applications." *CZECH TECHNICAL UNIVERSITY, DEPARTMENT OF*. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.24.419>
- Osterwalder, A, and Y Pigneur (2004). "An ontology for e-business models." *Value creation from e-business models*. http://inforge.unil.ch/yp/Pub/Final_BookChapter.pdf
- Ovchinnikova, E. (2012). "Natural Language Understanding and World Knowledge." In *Integration of World Knowledge for Natural Language Understanding* (pp. 15-37). Atlantis Press. http://link.springer.com/chapter/10.2991/978-94-91216-53-4_2
- Pace, Dale K. (2000). "Conceptual model development for C4ISR simulations." In *Proceedings of the 5th international command and control research and technology symposium*, Laurel, Maryland. Washington, D.C. USA., p. 24-26.
- Paltoglou, Georgios, and Mike Thelwall (2012). "Twitter, MySpace, Digg." *ACM Transactions on Intelligent Systems and Technology* 3(4): 1-19. <http://dl.acm.org/citation.cfm?id=2337542.2337551>
- Parry, David (2006). 1 Capturing Intelligence *Fuzzy Logic and the Semantic Web*. Elsevier. <http://www.sciencedirect.com/science/article/pii/S1574957606800043>
- Perez-Carballo, Jose, and Tomek Strzalkowski (2000). "Natural language information retrieval: progress report." *Information Processing & Management* 36(1): 155-178. <http://www.sciencedirect.com/science/article/pii/S0306457399000497>
- Perzanowski, D., A.C. Schultz, and W. Adams (1998). "Integrating natural language and gesture in a robotics domain." In *Proceedings of the 1998 IEEE International Symposium on Intelligent Control (ISIC) held jointly with IEEE International Symposium on Computational Intelligence in Robotics and Automation (CIRA) Intelligent Systems and Semiotics (ISAS) (Cat. No.98CH36262)*, IEEE, p. 247-252. <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=713669>
- Pisanelli, DM, A Gangemi, and G Steve (2002). "Ontologies and information systems: the marriage of the century?" *New Trends in Software Methodologies, Tools and Techniques*, 125-133. <http://books.google.es/books?hl=es&lr=&id=w4TFm5Dr3dIC&oi=fnd&pg=PA125&dq=Ontologies+and+Information+Systems:+the+Marriage+of+the+Century?&ots=P46LNn5qSz&sig=OIIdj0CwEcgAxw9vA77bDSOBuVk>
-

-
- Pitt, L, and R van der Merwe (2006). "Global alliance networks: A comparison of biotech SMEs in Sweden and Australia." *Industrial Marketing Management*, 35(5), 600-610.
<http://www.sciencedirect.com/science/article/pii/S0019850106000496>
- Qin, H, X Ma, T Herawan, and JM Zain (2011). "An adjustable approach to interval-valued intuitionistic fuzzy soft sets based decision making." *Intelligent Information and Database*
http://link.springer.com/chapter/10.1007/978-3-642-20042-7_9
- Qin, Li, and Vijayalakshmi Atluri (2009). "Evaluating the validity of data instances against ontology evolution over the Semantic Web." *Information and Software Technology* 51(1): 83–97.
<http://www.sciencedirect.com/science/article/pii/S0950584908000244>
- Raimond, Y, P Sinclair, NJ Humfrey, and M Smethurst (2009). "BBC programmes ontology."
<http://www.bbc.co.uk/ontologies/programmes/2009-09-07.shtml>
- Rana, Chhavi, and Sanjay Kumar Jain (2012). "A study of the dynamic features of recommender systems." *Artificial Intelligence Review*. <http://link.springer.com/10.1007/s10462-012-9359-6>
- Redondo, R. G. (2007). "El exceso de información en Internet convierte simples necesidades en retos. Entrevista a Jeff Barr."
http://scholar.google.es/scholar?q=Entrevista+a+Jeff+Barr&btnG=&hl=es&as_sdt=0,5#0
- Reilly, James, Kevin McCarthy, Lorraine McGinty, and Barry Smyth (2004). "Dynamic Critiquing." (03): 763–777.
- Resnick, Paul, N Iacovou, and M Suchak (1994). "GroupLens: an open architecture for collaborative filtering of netnews." *Proceedings of the 1994 ACM conference on Computer supported cooperative work* (pp. 175-186). ACM.: 175–186. <http://dl.acm.org/citation.cfm?id=192905>
- Resnick, Paul, Neophytos Iacovou, Mitesh Suchak, Peter Bergstrom, and John Riedl (1994). "GroupLens." In *Proceedings of the 1994 ACM conference on Computer supported cooperative work - CSCW '94*, New York, New York, USA: ACM Press, p. 175–186.
<http://dl.acm.org/citation.cfm?id=192844.192905>
- Resnick, Paul, and Hal R. Varian (1997). "Recommender systems." *Communications of the ACM* 40(3): 56–58. <http://dl.acm.org/citation.cfm?id=245108.245121>
- Ricci, F, L Rokach, and B Shapira (2011). "Introduction to recommender systems handbook."
http://link.springer.com/chapter/10.1007/978-0-387-85820-3_1
-

-
- Rich, E. (1979). "Building and exploiting user models." *Proceedings of the 6th international joint conference on Artificial intelligence-Volume 2* (pp. 720-722). Morgan Kaufmann Publishers Inc.
<http://dl.acm.org/citation.cfm?id=1623079>
- Rich, Elaine. (1983). "Users are individuals: individualizing user models." *International Journal of Man-Machine Studies* 18(3): 199-214.
<http://www.sciencedirect.com/science/article/pii/S0020737383800078>
- Richards, T, and L Richards (1994). "Using computers in qualitative research." *Methods*.
[http://www.educ.ttu.edu/uploadedFiles/personnel-folder/lee-duemer/epsy-6304/documents/Using computers in qualitative research.pdf](http://www.educ.ttu.edu/uploadedFiles/personnel-folder/lee-duemer/epsy-6304/documents/Using%20computers%20in%20qualitative%20research.pdf)
- Robins, Garry, Tom Snijders, Peng Wang, Mark Handcock, and Philippa Pattison (2007). "Recent developments in exponential random graph (p^*) models for social networks." *Social Networks* 29(2): 192-215. <http://www.sciencedirect.com/science/article/pii/S0378873306000384>
- Robinson, Stewart (2006). "Conceptual Modeling for Simulation: Issues and Research Requirements." In *Proceedings of the 38th Conference on Winter Simulation*, WSC '06, Winter Simulation Conference, p. 792-800. <http://dl.acm.org/citation.cfm?id=1218112.1218259>.
- Rodríguez-García, M.Á., Rodríguez-González, A., Colomo-Palacios, R., Valencia-García, R., Gómez-Berbís, J. M., & García-Sánchez, F. (2014). Using Data Crawlers and Semantic Web to Build Financial XBRL Data Generators: The SONAR Extension Approach. *The Scientific World Journal*, 2014. <http://www.hindawi.com/journals/tswj/2014/506740/abs/>
- Rodríguez-González, A., Torres-Niño, J., Jimenez-Domingo, E., Gomez-Berbis, M. J., & Alor-Hernandez, G. (2012a). "AKNOBAS: A knowledge-based segmentation recommender system based on intelligent data mining techniques." *Computer and Information Systems*, 9(2), 713-740. <http://www.doiserbia.nb.rs/Article.aspx?ID=1820-02141200008R>
- Rodríguez-González, A., García-Crespo, Á., Colomo-Palacios, R., Gómez-Berbís, J. M., & Jiménez-Domingo, E. (2012b). "Using ontologies in drug prescription: the semMed approach." *Intelligence Methods and Systems Advancements for Knowledge-Based Business*, 247. https://books.google.es/books?hl=es&lr=&id=xK2eBQAAQBAJ&oi=fnd&pg=PA247&dq=Using+Ontologies+in+Drug+Prescription:+The+SemMed+Approach.+&ots=LpYhQH-NHD&sig=Zo-5xklh-ebfKTqbN_GLope0zyE
- Rodríguez-González, Alejandro, Ricardo Colomo-Palacios, Fernando Guldres-Iglesias, JuanMiguel Gómez-Berbís, and Angel García-Crespo (2012). "FAST: Fundamental Analysis Support for Financial Statements. Using semantics for trading recommendations." *Information Systems Frontiers* 14(5): 999-1017. <http://dx.doi.org/10.1007/s10796-011-9321-1>.
-

-
- Rodriguez-Gonzalez, Alejandro, Fernando Guldris-Iglesias, Ricardo Colomo-Palacios, Giner Alor-Hernandez, and Ruben Posada-Gomez (2010). "Improving N calculation of the RSI financial indicator using neural networks." In *2010 2nd IEEE International Conference on Information and Financial Engineering*, IEEE, p. 49–53.
<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5609252>
- Rodríguez-González, Alejandro, Javier Torres-Niño, Gandhi Hernández-Chan, Enrique Jiménez-Domingo, and Jose Maria Alvarez-Rodríguez (2012). "Using agents to parallelize a medical reasoning system based on ontologies and description logics as an application case." *Expert Systems with Applications* 39(18): 13085–13092.
<http://www.sciencedirect.com/science/article/pii/S0957417412008093>
- Rodriguez-Gonzalez, A., Hernandez-Chan, G., Colomo-Palacios, R., Miguel Gomez-Berbis, J., Garcia-Crespo, A., Alor-Hernandez, G., & Valencia-Garcia, R. (2012). "Towards an ontology to support semantics enabled diagnostic decision support systems." *Current ...* 7: 234–245.
<http://www.ingentaconnect.com/content/ben/cbio/2012/00000007/00000003/art00002>
- Rodríguez-González, A., Labra-Gayo, J. E., Colomo-Palacios, R., Mayer, M. A., Gómez-Berbis, J. M., & García-Crespo, A. (2012). "SeDeLo: using semantics and description logics to support aided clinical diagnosis." *Journal of medical ...* <http://link.springer.com/article/10.1007/s10916-011-9714-1>
- Rosaci, Domenico, and Giuseppe M.L. Sarné (2010). "EFFICIENT PERSONALIZATION OF E-LEARNING ACTIVITIES USING A MULTI-DEVICE DECENTRALIZED RECOMMENDER SYSTEM." *Computational Intelligence* 26(2): 121–141. <http://doi.wiley.com/10.1111/j.1467-8640.2009.00343.x>
- Ruiz-Montiel, M, and JF Aldana-Montes (2009). "Semantically enhanced recommender systems." ... *Move to Meaningful Internet Systems* http://link.springer.com/chapter/10.1007/978-3-642-05290-3_74
- Saha Roy, Rishiraj, Rahul Katore, Niloy Ganguly, Srivatsan Laxman, and Monojit Choudhury (2014). "Discovering and understanding word level user intent in Web search queries." *Web Semantics: Science, Services and Agents on the World Wide Web* 30: 22–38.
<http://www.sciencedirect.com/science/article/pii/S1570826814000584>
- Sanchez, E. (2006). "Fuzzy logic and the semantic web." <https://books.google.es/books?hl=es&lr=&id=Cidej8b4ESIC&oi=fnd&pg=PP2&dq=A+fuzzy+description+logic+for+the+semantic+web,+in:+E.+Sanch&ots=mv9583ebL7&sig=v9piJKxXvLCYiQnW3kY83xukc2g>
-

-
- Schafer, JB, JA Konstan, and J Riedl (2001). "E-commerce recommendation applications." *Applications of Data Mining to Electronic Commerce* (pp. 115-153). Springer US.
http://link.springer.com/chapter/10.1007/978-1-4615-1627-9_6
- Schiaffino, Silvia, and Analía Amandi (2009). "Building an expert travel agent as a software agent." *Expert Systems with Applications* 36(2): 1291–1299.
<http://www.sciencedirect.com/science/article/pii/S0957417407005763>
- Schlenoff, C, and E Prestes (2012). "An IEEE standard ontology for robotics and automation." In *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on* (pp. 1337-1342). IEEE..
http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6385518
- Schockaert, S, P Victor, and GJ Houben (2006). "Reflections on Modelling Vagueness in Description Logics." *URSW*. <http://www.cwi.ugent.be/Chris/sschockaert.pdf>
- Serrano-Guerrero, Jesus, Enrique Herrera-Viedma, Jose A. Olivas, Andres Cerezo, and Francisco P. Romero (2011). "A google wave-based fuzzy recommender system to disseminate information in University Digital Libraries 2.0." *Information Sciences* 181(9): 1503–1516.
<http://www.sciencedirect.com/science/article/pii/S0020025511000223>
- Shadbolt, N, W Hall, and T Berners-Lee (2006). "The semantic web revisited." *Intelligent Systems, IEEE*. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1637364
- Shah, Urvi, Tim Finin, and Anupam Joshi (2002). "Information retrieval on the semantic web." In *Proceedings of the eleventh international conference on Information and knowledge management - CIKM '02*, New York, New York, USA: ACM Press, p. 461.
<http://dl.acm.org/citation.cfm?id=584792.584868>
- Shani, Guy, and Asela Gunawardana (2011). "Evaluating recommendation systems." *Recommender systems handbook*: 1–41. http://link.springer.com/chapter/10.1007/978-0-387-85820-3_8
- Shapira, B, L Rokach, and S Freilikhman (2013). "Facebook single and cross domain data for recommendation systems." *User Modeling and User-Adapted Interaction*, 23(2-3), 211-247.
<http://link.springer.com/article/10.1007/s11257-012-9128-x>
- Sheth, BD. (1994). "A learning approach to personalized information filtering." <http://pubs.media.mit.edu/pubs/papers/sheth-thesis.pdf>
- Simonsson, Márten, Robert Lagerström, and Pontus Johnson (2008). "A Bayesian network for IT governance performance prediction." In *Proceedings of the 10th international conference on*
-

-
- Electronic commerce - ICEC '08*, New York, New York, USA: ACM Press, p. 1.
<http://dl.acm.org/citation.cfm?id=1409540.1409542>
- Simperl, E. (2009). "Reusing ontologies on the Semantic Web: A feasibility study." *Data & Knowledge Engineering*. <http://www.sciencedirect.com/science/article/pii/S0169023X0900007X>
- Sizov, S. (2010). "Geofolk: latent spatial semantics in web 2.0 social media." In *Proceedings of the third ACM international conference on Web search and data mining* (pp. 281-290). ACM.
<http://dl.acm.org/citation.cfm?id=1718522>
- Smith, B, W Ceusters, and B Klagges (2005). "Relations in biomedical ontologies." *Genome biology*, 6(5), R46. <http://www.biomedcentral.com/1465-6906/6/R46>
- Staab, Steffen, and Rudi Studer (2004). *Handbook on Ontologies*. Berlin, Heidelberg: Springer Berlin Heidelberg. <http://link.springer.com/10.1007/978-3-540-24750-0>
- Stavrianou, Anna, and Caroline Brun (2013). "EXPERT RECOMMENDATIONS BASED ON OPINION MINING OF USER-GENERATED PRODUCT REVIEWS." *Computational Intelligence*: n/a–n/a. <http://doi.wiley.com/10.1111/coin.12021>
- Stephen, AT, and O Toubia (2010). "Deriving value from social commerce networks." *Journal of marketing research*. <http://journals.ama.org/doi/abs/10.1509/jmkr.47.2.215>
- Steve, G, A Gangemi, and DM Pisanelli (1998). "Integrating Medical Terminologies with ONIONS Methodology." *Information Modelling and Knowledge Bases IX*, 9, 1.
<http://books.google.es/books?hl=es&lr=&id=c0y7VnzLrx8C&oi=fnd&pg=PA1&dq=Integrating+Medical+Terminologies+with+ONIONS+Methodology.+&ots=CA48y03R7d&sig=fgKjh1UUabRfvgJ-nPvWOMXe0qY>
- Stodolsky, D.S. (1995). "Consensus journals: invitational journals based upon peer review." *The Information Society*. <http://www.tandfonline.com/doi/abs/10.1080/01972243.1995.9960199>
- Stoilos, G, G Stamou, and JZ Pan (2010). "Fuzzy extensions of OWL: Logical properties and reduction to fuzzy description logics." *International Journal of Approximate Reasoning*, 51(6), 656-679. <http://www.sciencedirect.com/science/article/pii/S0888613X10000125>
- Stoilos, G, GB Stamou, V Tzouvaras, JZ Pan, and I Horrocks (2005). "Fuzzy OWL: Uncertainty and the Semantic Web." *OWLED*. <http://www.image.ntua.gr/papers/398.pdf>
- Straccia, U. (2006). "A fuzzy description logic for the semantic web." *Capturing Intelligence*. <http://www.sciencedirect.com/science/article/pii/S1574957606800067>
-

-
- Studer, R, S Grimm, and A Abecker (2007). "Semantic web services: concepts, technologies, and applications." http://books.google.es/books?hl=es&lr=&id=wglTWruS3UC&oi=fnd&pg=PA1&dq=technology+of+the+web+semantic&ots=E4lvB20Yc3&sig=GDpxzFv_REP13IXM6to2OGb7u-0
- Studer, Rudi, V. Richard Benjamins, and Dieter Fensel (1998). "Knowledge engineering: Principles and methods." *Data & Knowledge Engineering* 25(1-2): 161-197. <http://www.sciencedirect.com/science/article/pii/S0169023X97000566>.
- Tamayo, Mario (2004). *El proceso de la investigación científica*. 4th ed. Editorial Limusa.
- Tani, MYK, and A Lablack (2014). "An ontology based approach for inferring multiple object events in surveillance domain." In *Science and Information Conference (SAI)*, 2014 (pp. 404-409). IEEE. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6918219
- Teller, J, and AK Keita (2007). "Urban Ontologies for an improved communication in urban civil engineering projects. Presentation of the COST Urban Civil Engineering Action C21". *TOWNTOLGY". Cybergeog: European Journal of Geography*. <http://cybergeog.revues.org/8322?lang=fr&file=1>
- Terveen, L, W Hill, and B Amento (1997). "PHOAKS: A system for sharing recommendations." *Communications of the ACM*, 40(3), 59-62. <http://dl.acm.org/citation.cfm?id=245122>
- Thelwall, M (2011). "Sentiment in Twitter events." *Journal of the Society for Information Science and Technology*, 62(2), 406-418. <http://onlinelibrary.wiley.com/doi/10.1002/asi.21462/full>
- Thompson, RL, CA Higgins, and JM Howell (1991). "Personal computing: toward a conceptual model of utilization." *MIS quarterly*. <http://www.jstor.org/stable/249443>
- Tiroshi, Amit, Tsvi Kuflik, Judy Kay, and Bob Kummerfeld (2012). "Recommender Systems and the Social Web."
- Trajkova, J, and S Gauch (2004). "Improving Ontology-Based User Profiles." *RIAQ*. [http://eolo.cps.unizar.es/docencia/MasterUPV/Articulos/Improving Ontology-Based User Profiles.pdf](http://eolo.cps.unizar.es/docencia/MasterUPV/Articulos/Improving%20Ontology-Based%20User%20Profiles.pdf)
- Trinkunas, J, and O Vasilecas (2015). "A graph oriented model for ontology transformation into conceptual data model." *Information Technology And Control*. <http://www.kalbos.ktu.lt/index.php/ITC/article/view/11832>
- Tsatsou, Dorothea, Vasileios Mezaris, and Ioannis Kompatsiaris (2012). "Semantic Personalisation in Networked Media: Determining the Background Knowledge." In *2012 Seventh International*
-

-
- Workshop on Semantic and Social Media Adaptation and Personalization*, IEEE, p. 101–106.
<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6406825>
- Tudorache, T, J Vendetti, and NF Noy (2008). “Web-Protege: A Lightweight OWL Ontology Editor for the Web.” *OWLED*. https://owl1-1.googlecode.com/svn-history/r642/trunk/www.webont.org/owled/2008/papers/owled2008eu_submission_40.pdf
- Tumasjan, A, TO Sprenger, PG Sandner, and IM Welp (2010). “Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment.” *ICWSM*.
<http://www.aaai.org/ocs/index.php/ICWSM/ICWSM10/paper/viewFile/1441/1852>
- Turing, AM. (1950). “Computing machinery and intelligence.” *Mind*.
<http://www.jstor.org/stable/2251299>
- Turing, A. M. (2009). “Computing machinery and intelligence.” *Parsing the Turing Test*.
http://link.springer.com/chapter/10.1007/978-1-4020-6710-5_3
- Valdéz, ER Núñez (2012). “Sistemas de recomendación de contenidos para libros inteligentes.”
<http://digibuo.uniovi.es/dspace/handle/10651/13126>
- Valencia-García, R. (2014). “Special issue on Systems Development by Means of Semantic Technologies.” *Science of Computer Programming*, 95, 1-2.
<http://www.sciencedirect.com/science/article/pii/S0167642314001749>
- Valencia-García, R, and F García-Sánchez (2009). “INNOVA3. 0 semantic platform for open innovation in financial domains.” <http://www.wseas.us/e-library/conferences/2009/tenerife/EACT-ISP/EACT-ISP-23.pdf>
- Vallet, D, M Fernández, and P Castells (2005). “An ontology-based information retrieval model.” ... *Web: Research and Applications*. http://link.springer.com/chapter/10.1007/11431053_31
- Varelas, G, and E Voutsakis (2005). “Semantic similarity methods in wordNet and their application to information retrieval on the web.” In *Proceedings of the 7th annual ACM international workshop on Web information and data management* (pp. 10-16). ACM.
<http://dl.acm.org/citation.cfm?id=1097051>
- Vásquez, Augusto Cortez, Hugo Vega huerta, Jaime Pariona Quispe, and Ana Maria Huayna (2009). “Procesamiento de lenguaje natural.” *Revista de investigación de Sistemas e Informática* 6(2): 45–54. <http://revistasinvestigacion.unmsm.edu.pe/index.php/sistem/article/view/5923>
-

-
- Vásquez-Ramírez, R. (2014). "Athena: A hybrid management system for multi-device educational content." *Computer Applications in Engineering Education*, 22(4), 750-763. <http://onlinelibrary.wiley.com/doi/10.1002/cae.21567/full>
- Veil, Shari R., Tara Buehner, and Michael J. Palenchar (2011). "A Work-In-Process Literature Review: Incorporating Social Media in Risk and Crisis Communication." *Journal of Contingencies and Crisis Management* 19(2): 110–122. <http://doi.wiley.com/10.1111/j.1468-5973.2011.00639.x>
- Vela, Mihaela, and Thierry Declerck (2009). "Concept and relation extraction in the finance domain." : 346–350. <http://dl.acm.org/citation.cfm?id=1693756.1693801>
- Vergara, Margarita, and Salvador Mondragón (2014). "CONTRIBUCIONES A LA TAXONOMÍA DE LA SEMÁNTICA DE PRODUCTOS." http://www.researchgate.net/publication/239532252_CONTRIBUCIONES_A_LA_TAXONOMIA_DE_LA_SEMNTICA_DE_PRODUCTOS
- Verma, S, S Vieweg, WJ Corvey, and L Palen (2011). "Natural Language Processing to the Rescue? Extracting ' Situational Awareness' Tweets During Mass Emergency." *ICWSM*. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.371.9621&rep=rep1&type=pdf>
- Viinikkala, M. (2003). "Ontology in Information Systems." *Disponible en*. [http://scholar.google.es/scholar?q=Ontology+in+Information+Systems".+&btnG=&hl=es&as_sdt=0,5#1](http://scholar.google.es/scholar?q=Ontology+in+Information+Systems)
- Voorhees, Ellen M. (1999). "Natural Language Processing and Information Retrieval." *Information extraction*: 32–48. http://dx.doi.org/10.1007/3-540-48089-7_3.
- Vozalis, M. and K Margaritis (2007). "Using SVD and demographic data for the enhancement of generalized Collaborative Filtering." *Information Sciences* 177(15): 3017–3037. <http://www.sciencedirect.com/science/article/pii/S0020025507001223>
- Vysniauskas, E, and L Nemuraite (2015). "Transforming ontology representation from OWL to relational database." *Information technology and control*. <http://www.inzeko.ktu.lt/index.php/ITC/article/view/11779>
- Wagner, WP, J Otto, and QB Chung (2002). "Knowledge acquisition for expert systems in accounting and financial problem domains." *Knowledge-Based Systems*. <http://www.sciencedirect.com/science/article/pii/S0950705102000266>
-

-
- Wang, Rui-Qin, and Fan-Sheng Kong (2007). "Semantic-Enhanced Personalized Recommender System." In *2007 International Conference on Machine Learning and Cybernetics*, IEEE, p. 4069–4074. <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4370858>
- Wang, W, D Zhao, H Luo, and X Wang (2013). "Mining User Interests in Web Logs of an Online News Service Based on Memory Model." In *Networking, Architecture and Storage (NAS), 2013 IEEE Eighth International Conference on* (pp. 151-155). IEEE.. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6665357
- Weaver, AC, and BB Morrison (2008). "Social networking." *Computer*. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4454412
- Weinberg, Bruce D., and Ekin Pehlivan (2011). "Social spending: Managing the social media mix." *Business Horizons* 54(3): 275–282. <http://www.sciencedirect.com/science/article/pii/S0007681311000255>
- Weng, J, and BS Lee (2011). "Event Detection in Twitter." *ICWSM*. <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM11/paper/download/2767/3299>
- Widyantoro, D, and J Yen (2001). "Incorporating fuzzy ontology of term relations in a search engine." *Proceedings of the BISC Int. Workshop on Fuzzy Logic and the Internet* (pp. 155-160). https://scholar.google.es/scholar?q=%22Incorporating+Fuzzy+Ontology+of+Term+Relations+in+a+Search+Engine&btnG=&hl=es&as_sdt=0%2C5#0
- Wolfe, Alvin W. (1997). "Social Network Analysis: Methods and Applications." *American Ethnologist* 24(1): 219–220. <http://doi.wiley.com/10.1525/ae.1997.24.1.219>
- Xiang, Liang, and Qing Yang (2009a). "Time-Dependent Models in Collaborative Filtering Based Recommender System." *2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology*. 450–457. <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5286031>
- Xiang, L., & Yang, Q. (2009b). "Time-Dependent Models in Collaborative Filtering Based Recommender System." *2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology*. 450–457. <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5286031>
- Xie, G, R Xiong, and I Church (1998). "Comparison of kinetics, neural network and fuzzy logic in modelling texture changes of dry peas in long time cooking." *LWT-Food Science and Technology*. <http://www.sciencedirect.com/science/article/pii/S0023643898904169>
-

-
- Ye, J, and L Coyle (2007). "Ontology-based models in pervasive computing systems." *The Knowledge* http://journals.cambridge.org/abstract_S0269888907001208
- Zadeh, L.A. (1965). "Fuzzy sets." *Information and Control* 8(3): 338–353. <http://www.sciencedirect.com/science/article/pii/S001999586590241X>
- Zadeh, L. A. (1983). "The role of fuzzy logic in the management of uncertainty in expert systems." *Fuzzy Sets and Systems* 11(1-3): 197–198. <http://www.sciencedirect.com/science/article/pii/S0165011483800815>
- Zavitsanos, E, and G Paliouras (2007). "Discovering subsumption hierarchies of ontology concepts from text corpora." In *Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence* (pp. 402-408). IEEE Computer Society. <http://dl.acm.org/citation.cfm?id=1331880>
- Zeng, D, H Chen, R Lusch, and SH Li (2010). "Social media analytics and intelligence." *Intelligent Systems, IEEE*. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5678581
- Zha, XF, HJ Du, and JH Qiu (2001). "Knowledge-based approach and system for assembly-oriented design, Part II: the system implementation." *Engineering Applications of Artificial Intelligence*. <http://www.sciencedirect.com/science/article/pii/S0952197600000610>
- Zhang, Z, C Zhang, and S San Ong (2000). "Building an ontology for financial investment." In *Intelligent Data Engineering and Automated Learning—IDEAL 2000. Data Mining, Financial Engineering, and Intelligent Agents* (pp. 308-313). Springer Berlin Heidelberg. 5. http://link.springer.com/chapter/10.1007/3-540-44491-2_44
- Zhong, N, J Liu, and Y Yao (2002). "In search of the wisdom web." *Computer*. <http://www.computer.org/csdl/mags/co/2002/11/ry027.pdf>
- Zhou, Xujuan, Yue Xu, Yuefeng Li, Audun Josang, and Clive Cox (2011). "The state-of-the-art in personalized recommender systems for social networking." *Artificial Intelligence Review* 37(2): 119–132. <http://link.springer.com/10.1007/s10462-011-9222-1>
-

